Summer 1-1-2016

Microflora of the equine gut and its ramifications on the development of laminitis; A comparison of fecal and cecal diversity and Illumina and Roche 454 sequencers

Matthew John Jevit

MICROFLORA OF THE EQUINE GUT AND ITS RAMIFICATIONS ON THE DEVELOPMENT OF LAMINITIS; A COMPARISON OF FECAL AND CECAL DIVERSITY AND ILLUMINA AND ROCHE 454 SEQUENCERS

A Thesis

Submitted to the Bayer School of Natural and Environmental Science

Duquesne University

In partial fulfillment of the requirements for

the degree of Master of Science

By

Matthew J. Jevit

August 2016

MICROFLORA OF THE EQUINE GUT AND ITS RAMIFICATIONS ON THE DEVELOPMENT OF LAMINITIS; A COMPARISON OF FECAL AND CECAL DIVERSITY AND ILLUMINA AND ROCHE 454 SEQUENCERS

By

Matthew J. Jevit

Approved June 17th, 2016

Jan Janecka PhD Assistant Professor of Biology (Committee Chair)	Nancy Trun PhD Associate Professor of Biology (Committee Member)
Becky Morrow DVM Assistant Professor of Biology (Committee Member)	
Philip Reader PhD Dean, Bayer School of Natural and Environmental Science	John Stolz PhD Chair, Center for Environmental Research and Education Professor of Biology

ABSTRACT

MICROFLORA OF THE EQUINE GUT AND ITS RAMIFICATIONS ON THE DEVELOPMENT OF LAMINITIS; A COMPARISON OF FECAL AND CECAL DIVERSITY AND ILLUMINA AND ROCHE 454 SEQUENCERS

By

Matthew J. Jevit

August 2016

Dissertation supervised by Jan Janecka PhD

Laminitis is characterized by the separation of the phalanx and the hoof wall. It can be induced in horses by ingesting high amounts of non-structural carbohydrates (NSC), which changes the hindgut microflora. However, fecal bacteria may not be representative of the cecum. In addition, in horses results from more recent sequencers (Illumina) have never been compared to previously used sequencers (454). To determine if there are functional differences alpha and beta-diversity, core biomes, and shifts in hindgut bacteria in response to NSC were compared between fecal and cecal communities and the MiSeq and 454 method. The results suggest that MiSeq is superior to the 454 due to greater number of reads per cost. The method had a greater effect on the diversity than the sample origin. Fecal microflora exhibited more substantial shifts than the cecum. It is hypothesized this is due to the downstream migration of lactic acid and VFAs.

DEDICATION

This thesis is dedicated to the memory of Dr. Josie Coverdale. Dr. Coverdale was dedicated to the research and wellbeing of horse and understanding of equine nutrition, as well as the personal and professional development of her students. Dr. Coverdale was one of the primary driving forces behind this project but was unfortunately unable to see its completion. She was a wonderful colleague, mentor, and friend to those that knew her and is greatly missed.

ACKNOWLEDGEMENT

I would like to thank my chair, Dr. Jan Janecka, as well as my committee members, Dr. Nancy Trun, and Dr. Becky Morrow, for their guidance and support throughout this project. In addition, I would like to thank Dr. Josie Coverdale and Christine Warzecha at Texas A&M University for their previous work and support of this project. Finally, I would like to thank Dr. John Stolz and the Center for Environmental Research and Education at Duquesne University for their support during the duration of this thesis. The Morris Animal Foundation, Duquesne University and Texas A&M University provided funding for this project.

TABLE OF CONTENTS

Page

Abstract iv
Dedication v
Acknowledgement vi
List of Tables ix
List of Figures x

Chapter 1. A Review of Current Literature on Equine Laminitis 1
Introduction 1
Pathology of Laminitis 2
The Equine Digestive System 5
Microbiology of the GI Tract of Horses 8
Pathophysiology of Laminitis Related to the Digestive Tract 11
Characterization of Bacterial Communities 15
Diversity Indexes 15
Alpha diversity 16
Beta Diversity 18
Justification of Research 19
Chapter 2. Material and Methods 22
Study animals and Experimental Treatments 22
Samples 23
Isolation of Microbial DNA 24
Amplification of DNA 25
Data Analysis 26
Chapter 3. Comparison of Two Sequencing Strategies; the Roche 454 FLX and the Illumina MiSeq 29
Comparison of the reads produced by each platform 29
Comparison of alpha diversity between the two sequencing 30
Comparison of the Beta Diversity 40
Comparison of Core Microbiomes 42
Identification of Shifts in GI Flora by the Illumina MiSeq and the Roche 454 FLX 46
Chapter 4. Comparison of Microflora in Fecal and Cecal Samples using the Roche 454 FLX High Throughput Sequencer 50

Reads generated in each sample set 50
Comparison of alpha diversity between fecal versus cecal 51
Comparison of the Beta Diversity 57
Identification of Shifts in GI Flora in Fecal and Cecal Samples 61
Chapter 5. Discussion and Conclusions 66
Comparison of Fecal Samples Sequenced with the Illumina MiSeq and the Roche 454 FLX 66
Comparing the Ability of Each Protocol to Identify OTUs 66
Comparing the Core Microbiomes Identified by Either Protocol 68
Comparing the Apparent Shifts in the Microflora in Response the High Carbohydrate Treatment 69
Conclusion 71
Comparison of Fecal and Cecal Bacteria 72
Comparing the Alpha Diversity of Each Sample Dataset 72
Comparison of the Beta Diversity between Fecal and Cecal Samples 73
Comparison of Core Biomes 73
Response to High Carbohydrate Treatment 74
Conclusions 76
Works Cited 77
Appendix 88

LIST OF TABLES

Table Number	Table	Page Number
1	Concentrations of nutritional components in feed	23
2	Comparison of alpha diversity statistics for the 454 fecal and MiSeq fecal	30
3	Pairwise comparison of fecal samples	34
4	Core OTUs identified in 95% of the Roche 454 fecal samples	43
5	Core OTUs identified in 95% of the Miseq fecal samples	43
6	Comparison of alpha diversity statistics for the 454 cecal and 454 fecal	50
7	Core OTUs identified in 95% of the Roche 454 fecal samples	60
8	Core OTUs identified in 95% of the Roche 454 cecal samples	61
9	Number of Reads identified in the MiSeq fecal and the 454 fecal samples	88
10	Weighted UniFrac values	90
11	Unweighted UniFrac values	90
12	GI microflora shifts observed in the MiSeq fecal samples	91
13	GI microflora shifts observed in the 454 fecal samples	91
14	Number of Reads identified in the 454 cecal and the 454 fecal samples	92
15	GI microflora shifts observed in the 454 fecal samples	94
16	GI microflora shifts observed in the 454 cecal samples	94
17	List of core OTUs identified in 95% of the Roche 454 fecal samples	95
18	List of core OTUs identified in 95% of the MiSeq fecal samples	95
19	List of core OTUs identified in 95% of the Roche 454 cecal samples	98

LIST OF FIGURES

Figure Number	Figure	Page Number
1	Illustration of the structure of the BM junction	3
2	MiSeq fecal and 454 Fecal Mean OTU richness	31
3	Number of OTUs versus Number of Reads	32
4	MiSeq fecal and 454 *chao1*estimates	33
5A	OTU richness rarefaction with 1,000 MiSeq reads	35
5B	*Chao1* rarefaction with 1,000 MiSeq reads	36
6A	OTU richness rarefaction with 10,000 MiSeq reads	37
6B	*Chao1* rarefaction with 10,000 MiSeq reads	38
7	*Chao1* rarefaction with 10,000 454 fecal reads	39
8	Standardized Simpson's D estimates of the MiSeq and the 454 fecal samples	40
9	PCoA of weighted *UniFrac* estimates	41
10	Bacterial families identified in the 454 fecal core biome	45
11	Bacterial families identified in the MiSeq fecal core biome	45
12	GI microflora shifts observed in the MiSeq fecal samples	47
13	GI microflora shifts observed in the 454 fecal samples	49
14	454 cecal and 454 fecal mean OTU richness	51
15	454 cecal and 454 fecal mean *chao1* estimates	52
16A	OTU richness rarefaction with 10,000 454 fecal reads	53
16B	*Chao1* rarefaction with 10,000 454 fecal reads	54
17A	OTU richness rarefaction with 10,000 454 cecal reads	55
17B	*Chao1* rarefaction with 10,000 454 cecal reads	56
18A	PCoA of weighted *UniFrac* estimates	58
18B	PCoA of weighted *UniFrac* estimates	59
19	GI microflora shifts observed in the 454 fecal samples	63
20	GI microflora shifts observed in the 454 cecal samples	65

Chapter 1. A Review of Current Literature on Equine Laminitis

Introduction

Laminitis, also known colloquially as founder, is a debilitating disease that affects the hooves of artiodactyls and perissodactyls. Laminitis occurs frequently in the domestic horse (*Equus ferus caballus*) (Perissodactyla) and the domestic cattle (*Bos taurus*) (Artiodactyla) (Nocek 1997; Pollitt 2004). Laminitis is a complex disease and it is characterized by the separation of the distal phalanx, referred to as the coffin bone in horses and the pedal bone in cattle, from the hoof wall (Pollitt 2004; Mungal et al. 2001). This separation can be extremely painful and debilitating. In severe cases, euthanasia is often the only humane treatment. As the coffin bone begins to separate, the anterior portion of the bone begins to rotate downwards into the sole of the hoof. This is due to the upward pressure placed on the posterior portion of the bone by the deep flexor tendon (White 2005). The subsequent downward pressure from the tip of the coffin bone is what causes a large amount the pain associated with laminitis.

The development of laminitis is not well-understood (Bailey et al. 2009). This may be because its development is associated with a number of different and unrelated factors (White 2005; Pollitt 2004). These factors can be local to the hoof, such as overuse, general wear and tear (e.g., walking on hard surfaces), or injury (White 2005). The development of laminitis is often connected with other conditions distant to the feet. These include colic, acidosis of the large intestine, retained placenta, bowel infarction, Potomac Horse Fever, and salmonella infection (Pollitt 2004; Kreuger et al. 1986; Hood et al. 1993).

Pathology of Laminitis

Laminitis is similar to equine colic in that both are combination of symptoms rather than a singular disease. Ultimately, laminitis is the separation of the coffin bone and the hoof wall, followed by the downward rotation of the coffin bone. This separation occurs at the junction of the dermis and the epidermis in the hoof. Finger like structures called primary laminae protrude from both the dermal and epidermal layers. The primary laminae have additional smaller laminae extending from them. These structures are responsible for the connection of the dermis and the epidermis. The secondary laminae of the dermas and the epidermis are connected to each other via a structure called the basement membrane (BM) that separates the two layers of skin (Pollitt 2004). The basement membrane is primarily made of protein fibers and provides attachment points for secondary laminae. Hemidesmosomes are small rivet like cellular structures found in the secondary laminae that provide attachment points to the BM (Pollitt 2004; French and Pollitt 2004) (Figure 1). When laminitis develops, laminar cells are observed to lose shape and become elongated. Additionally, the intercellular space between the secondary epidermal laminae and the BM increases with the severity of the condition. Blood vessels in the BM supplying the epidermis are retracted, causing localized ischemia (Pollitt 2004).

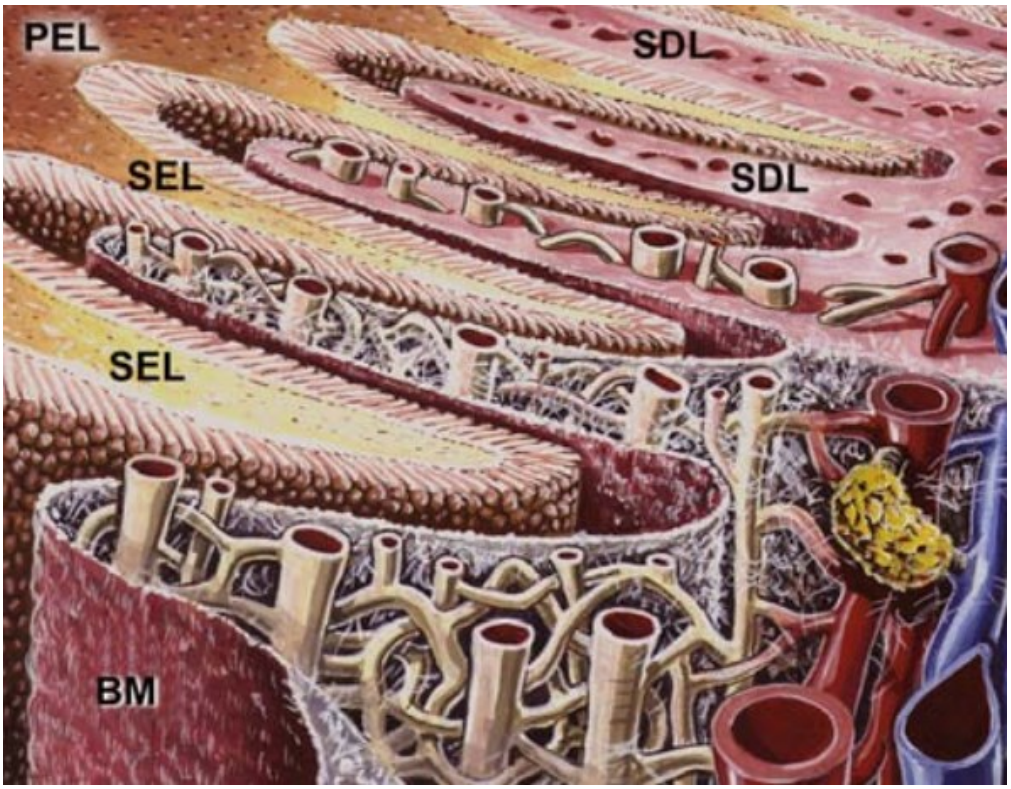

FIGURE 1. This diagram of the structure of the junction between the Dermal and the Epidermal Laminae was originally published in Pollitt (2008). The picture was designed by Christopher Pollitt and drawn by John McDougall. The basement membrane (BM) provides attachment points between the secondary dermal laminae (SDL) and the secondary epidermal laminae (SEL). The dermal tissue is ultimately connected to the coffin bone; whereas, the epidermis is connected to the hoof wall.

Laminitis, like colic, has several etiologies. These include causes that are both mechanical and physiological in origin. Mechanical causes include trauma to either an infected or contralateral limb. Trauma induced laminitis is often called road founder and is a result of overuse. For example, if the horse is forced to walk extensively on hard surfaces such as a concrete road, founder may occur (White 2005). Another mechanical etiology of laminitis is overuse of a specific leg due to injuries in other limbs. This is often caused by a fracture, or the development of arthritis or tenosynovitis in one or more limbs. Tenosynovitis is characterized by the inflammation of the synovium, which is a fluid filled sack that surrounds a tendon. These injuries cause the horse to favor a healthy limb leading to added stress and the development of laminitis in the healthy limb (White 2005; Van Epps et al. 2010; Toupadakis et al 2010).

In addition to mechanical causes of laminitis, several physiological causes have been identified. The development of laminitis is often associated with a diet that a horse has not evolved to consume. This includes the consumption of black walnut extract and overconsumption of feed high in nonstructural carbohydrates (NSC) such as starch and oligofructose (Al Jassim & Andrews 2012; Van Epps & Pollitt 2006; Longland & Byrd 2006). It is hypothesized that this type of laminitis is caused by abnormalities in the gastrointestinal microflora. These abnormal diets may lead to the dysregulation of hindgut microflora, proliferation of certain bacteria normally not abundant, and secretion of exotoxins/endotoxins thereby causing either sepsis, endotoxemia, or an immune response. These potential mechanisms then trigger laminitis. This model of development will be the primary focus and will be discussed at length later in this paper. In order to understand this model of development, the equine digestive system must first be understood.

The Equine Digestive System

Horses are hindgut fermenters, unlike ruminants such as cows, sheep, and goats, and therefore ferment structural carbohydrates (primarily cellulose) in the large intestine instead of the stomach. In addition, NSCs can also be fermented in the cecum, a specialized region of the large intestine (Al Jassim & Andrews 2009; Sjaastad et al. 2003; Weyenberg & Janssens 2005).

Horses have specialized regions of the digestive tract designed to digest and absorb primarily NSC, proteins, fatty acids, and lipids. In the oral cavity and the stomach, food is broken down mechanically through chewing or acidic degradation and if it is not adequately broken down it can become impacted in the intestines. Such intestinal impaction is associated with bowel infarctions, and is one of the causes of colic (Al Jassim & Andrews 2009). The horse has 3 pairs of salivary glands (parotid, mandibular, and submaxillary). The saliva that is excreted from these glands does not have any enzymatic properties and is solely for lubrication and to aid swallowing.

The stomach makes up 8-10% of the volume (7-15 litters) of the GI tract. The esophagus enters the stomach at a sharp angle making horses incapable of vomiting which causes health issues (Al Jassim & Andrews 2009). The stomach is divided into a proximal non-glandular region and a distal glandular region. The pH of the non-glandular portion of the stomach is ~5.4, whereas the glandular portion has a pH of ~1.8 (Al Jassim & Andrews 2009). Food that enters the stomach is not mixed with the contents. This ensures that the stomach acid produced in the lower portion of the stomach does not encounter the more sensitive upper portion (Al Jassim & Andrews 2009; Weyenberg & Janssens 2005). Chewing of food induces the secretion of saliva, which contains bicarbonate (HCO_3) that buffers the pH of the non-glandular stomach region.

Initial fermentation of NSCs occurs in the distal portion of the stomach and though some volatile fatty acids (VFA) are produced, the main product is lactic acid (Al Jassim & Andrews 2009).

The small intestine is comprised of three distinct regions; the duodenum, jejunum, and ileum. When chyme, acidic pulp from the stomach, enters the small intestine, the pancreas releases a bicarbonate secretion with digestive enzymes into the duodenum (Al Jassim & Andrews 2009). Additionally, the liver secretes bile directly into the duodenum. The bile and the bicarbonate secretions act to neutralize the hydrochloric acid from the stomach and create an environment with the optimal pH for pancreatic enzymes to breakdown lipids, NSCs, and proteins. In grass fed horses, the pH of the duodenum, jejunum, and ileum is ~ 6.32, ~7.10, and ~7.47, respectively (Al Jassim & Andrews 2009). Pancreatic fluid is secreted at an increasing rate, peaking 3-4 minutes after the chyme has entered the small intestine. The final concentration of pancreatic fluid in the duodenum is 3-4-fold higher than prior to feeding. In contrast, ruminants release pancreatic solution at a constant rate. Horses also secrete more pancreatic solution (10-12 liters/100 kg body weight) than ruminants (5 liters/100 kg body weight for a cow) (Al Jassim & Andrews 2009; Sjaastad et al. 2003). The bicarbonate and pancreatic enzymes are at a lower concentration in the pancreatic solution of horses than of ruminants. The continuous release of the pancreatic fluid and bile ensures the ingesta entering the large intestine is buffered so that the microbial communities in the cecum, the first portion of the large intestine, are not disrupted. The majority of proteins, lipids, and naturally occurring NSCs are digested and absorbed before the ingesta reaches the cecum (Al Jassim & Andrews 2009).

The large intestine of the horse is the largest digestive organ in terms of volume (64% total GI volume, 95-112 liters) and is the most significant for energy production. Equids do not

produce the necessary enzymes for the degradation of structural proteins such as cellulose. Bacteria in the GI tract ferment these structural carbohydrates into volatile fatty acids such as formic acid and acetic acid. These VFAs are then absorbed by the horse, metabolized into compounds including acetyl CoA, and subsequently used for energy production (Al Jassim & Andrews 2009; Cummings 1981; Miller and Wollin 1996). The large intestine is divided into three major regions; the cecum, colon and rectum, respectively. The walls are thick and comprised of longitudinal and circular muscle fibers. These walls lack villi but contain glands from which mucus is secreted (Al Jassim & Andrews 2009; Weyenberg & Janssens 2005).

The cecum (25-35 liters in volume) is responsible for most of the bacterial mediated fermentation (Al Jassim & Andrews 2009). The cecum has bands of muscles called tarnia, which close off portions of the cecum from each other. The closed off sections are called haustra; this serves to delay the time that it takes for chyme to move through the cecum as well as to mix the contents. After structural carbohydrates are fermented into VFAs, they enter the ascending colon where they are absorbed (Al Jassim & Andrews 2009). The ascending colon (150 liters in volume) is separated into two major portions; the dorsal and the ventral. The ascending colon is followed by the descending colon, which is connected, to the rectum. By the time the ingesta reaches the descending colon, the majority of the VFAs and other nutrients have been absorbed and the remaining material is defecated (Al Jassim & Andrews 2009). During digestion ingesta is present in the ascending and descending colon for the greatest amount of time (Al Jassim & Andrews 2009; Weyenberg & Janssens 2005).

The rumen of bovids has been used as a model for the hindgut fermentation in equids because of similarities in fermentation patterns with the cecum (Kern et al., 1974). One of the most common families of bacteria in the rumen and cecum is *Ruminococcaceae.* These species,

namely *Ruminococcus flavefaciens,* have been observed to ferment carbohydrates to succinic acid, but will also produce acetic and formic acids (Latham and Wollin 1971). Other VFAs, such as propionate and butyrate, are very common. Cellulose, the primary structural carbohydrate in the cell wall of plants, is the main source of fermentable material. Carbohydrates can be broken down via the Embden-Myerhoff-Parnas glycolytic pathway to pyruvate. Pyruvate is then metabolized into VFAs, H_2, CO_2, CH_2, and H_2O (Cummings 1981; Miller and Wollin 1996). Acetic acid (acetate) is produced by oxidative decarboxylation of pyruvate. Propionate is produced by the fixation of carbon dioxide forming succinate, which is then decarboxylized via the dicarboxylic acid pathway. Succinate can also be fed into the acrylate pathway resulting in the production of lactate and acrylate (Cummings 1981). Though lactate can be produced, it is often not observed at high frequency. This may be due to the fact that increased rates of fermentation lead to a decrease in gut pH, which often inhibits absorption and metabolism of lactose (Cummings 1981). Once VFAs are produced they are readily absorbed in the rumen and cecum. This most likely occurs as VFAs are protonated and allowed to diffuse in their non-ionized form via a gradient (Cummings 1981). Once in the blood stream or the epithelia cells of the gut wall, many of the VFAs can be converted into metabolic precursor, which can be fed into the Krebs cycle. For example, acetyl CoA can be produced from acetic and butyric acid and succinyl CoA can be produced from propionic acid.

Microbiology of the GI Tract of Horses

The bacteria of the GI tract have many functions related to the health of the horse (Al Jassim & Andrews 2009; Macpherson & Harris 2004). There are distinct bacterial communities throughout the digestive tract but the microbial communities are largest in the distal portions of the small intestine and the cecum (Mackie & Wilkins 1998). Previous studies exploring the

diversity of the equine hindgut microflora were carried out by cloning and Sanger sequencing 16S rRNA genes (Daly et al. (2001) and Milinovich et al. (2008)). Recently, studies utilizing next-generation sequencing methods have been published (Costa et al. 2012; Steelman et al. 2012; Dougal et al 2013; Moreau et al. 2014). The majority of the effort was focused on the fermenting bacteria of the cecum (Al Jassim & Anderws 2009; Al Jassim et al. 2005; Daly et al. 2001; Dougal et al. 2013; Kern et al. 1974; Mackie & Wilkins 1998). This was done in order to better understand digestion of food. More recently, with the realization that bacteria play a role in the development of diseases such as acidosis, colic, and laminitis, more studies have been conducted to determine how cecal bacteria contribute to the health of the horse (Al Jassim & Andrews 2009; Milinovich et al. 2006; Steelman et al. 2012).

The stomach, due to its low pH, has microflora composed of fast growing and acid producing bacteria dominated by lactic acid producers (Al Jassim et al. 2011). The cecum has the greatest amount of bacterial biomass with approximately $1.26X10^9$ bacteria per gram (wet weight) of ingesta. This is ~3-fold higher than the concentration of bacteria found in the small intestine. The duodenum, the jejunum and the ileum contain $2.9X10^6$, $2.9X10^7$, and $3.84X10^7$ bacteria per gram of ingesta, respectively (Al Jassim & Andrews 2009).

Cecal bacteria produce enzymes for fermentation of carbohydrates into VFAs. A few of the most commonly identified bacteria species are from the phyla *Firmicutes* (Biddle et al. 2013). This phylum is reported as accounting for 64%-70% of the total GI bacteria in the cecum depending on the publication (Shepard et al. 2012; Costa et al. 2011). Specifically, families such as *Ruminococcaceae,* and *Lachnospiraceae* are most common (Dougal et al. 2013). *Ruminococcaceae and Lachnospiraceae* are among the most abundant families in the GI tract of all mammals (Biddle et al. 2013). They are both responsible for the degradation of plant material

by breaking down structural carbohydrates such as cellulose (Daly et al. 2001). Horses are unable to digest plant matter without the aid of bacteria because equine digestive enzymes are incapable of severing the β-1, 4 glyosidic bonds that connect adjacent carbons in cellulose. This is a trait shared by all mammals (Vernon & Flint 1988). *Ruminococcaceae* secretes enzymes capable of breaking these bonds, and is capable of digesting a wide array of structural carbohydrates such as cellulose, hemicellulose, lignin, and pectin. It is important to note that although many species such as Ruminococcaceae can degrade structural carbohydrates, many prefer NSCs such as glucose (Biddle et al. 2013). *Ruminococcaceae,* and *Lachnospiraceae* are among the most biologically active bacteria in the mammalian gut. In humans, they make up 84% of the digestively active taxa and account for 30-50% of the Gram-positive bacteria (Biddle et al. 2013). The second most common bacterial phylum found in the cecum are the *Bactroidetes*, Gram-negative bacteria (Costa et al. 2011; Shepard et al, 2012). The most common *Bactroidetes* family is the *Prevotellaceae* (Dougal et al. 2013). Up to 33.3%-45% of all cecal bacteria, are anaerobes (Al Jassim & Andrews 2009; McCreery et al., 1971; Kern et al., 1973).

Digesta exhibits a short residence time in the small intestine. When excessive NSCs are consumed many will continue on to the cecum without being digested (Crawford et al. 2007, Al Jassim et al 2016). Digestively active Gram-positive bacteria such as *Ruminococcaceae* preferentially ferment these carbohydrates (Biddle et al. 2013). This leads to a rapid increase in the production of lactic acid and VFAs. In addition, the hindgut GI microflora composition shifts as the communities adapt to changing conditions (Grubb and Dehority 1975; Kern et al. 1974; Lloyd et al., 1956). The extent to which the hindgut microflora adapts is affected by a

number of different factors including magnitude and timing of the dietary change (Grubb and Dehority, 1975; Kern et al., 1974; Lloyd et al., 1956).

Pathophysiology of Laminitis Related to the Digestive Tract

The pathology of carbohydrate induced laminitis has been disputed (Hood et al. 1990; Pollitt & Davies 1998). An earlier hypothesis stated that ischemia led to the death of cellular structures in the hoof (Hood et al. 1990; Hood et al. 1993). Because of the close proximity of the coffin bone and the hoof wall, when swelling occurs the blood vessels in the hoof become pinched between the two structures and blood flow is reduced. This reduction in blood flow leads to the death of cellular structures that provide attachment to the BM. Hood et al. (1990) argued that equine laminitis was a similar to Raynaud's Phenomenon, which is characterized by reoccurring vasospasms in the hand and feet in response to cold temperatures. Like laminitis, Raynaud's Phenomenon can be associated with other disorders, most commonly systemic sclerosis (Block and Sequeira 2001). Observations in support of this hypothesis were that the hoof becomes warm during the acute phase of laminitis and that there is a rapid bounding pulse during the development of laminitis, both typical symptoms of swelling. However, the decreased blood flow was observed in the development phase of laminitis when no lameness is present (Hood et al. 1993). This hypothesis is generally no longer accepted.

The swelling in the hooves was proposed to originate from either sepsis or another source of systemic inflammation, most likely stemming from the GI microflora (Mungall et al 2001). One of the most common causes of equine laminitis is the over consumption of NSCs such as starch in feed used for racing and working horses. In addition, starch and simple sugars increase in pasture plants during the spring when they increase the amount of photosynthesis (Longland & Byrd 2006; Van Epps & Pollitt 2006; Garner et al. 1975; Kronfeld et al. 2006). It is

hypothesized that the excessive consumption of NSCs initiates a cascade in the hindgut, resulting changing microbial metabolism, proliferation of certain bacteria, and acidosis of the large intestine which in turn leads to a systemic inflammation (Pollitt 2004).

Digestion of NSCs in the upper GI tract is relatively inefficient because horses do not secrete adequate amounts of amylase to sustain a diet high in starch. On average, horses secrete only 8-10% of the amount of amylase secreted by a domestic pig (*Sus scrofa*) (Al Jassim & Andrews 2009). Horses have a slow digestive system and need to eat continuously in order to ensure that food is constantly moving through the large intestine to absorb energy (Al Jassim & Andrews 2009). Ingesta spends an inadequate among of time in the small intestine resulting in incomplete digestion and absorption of sugars and starch. Therefore, when excessive NSCs are consumed, the majority continue on to the cecum (Al Jassim & Andrews 2009). While in the large intestine, both nonstructural and structural carbohydrates are fermented into VFAs. Less complex carbohydrates ferment more quickly than the more complex carbohydrates (Edwards & Perret 2002; Rotger et al. 2005). Due to the increased rate at which NSCs can be fermented, an excess of VFAs and lactic acid are released into the lower intestine decreasing the pH. This can lead to a shift in the microflora as well as damage to the mucosal lining of the gut (Van Epps & Pollitt 2006; Longland & Byrd 2006; Pollitt 2004; Al Jassim & Andrews 2009). This hypothesis is supported by the observations that acidic diarrhea is exhibited in the development phase when laminitis is experimentally induced with either starch or oligofructose. In addition, increase in the abundance of fermenting bacteria, such as *Streptococcus* and *Lactobacillus* ssp., have been observed (Van den Berg et al. 2012; Pollitt 2004). Damage to the mucosal lining would lead to an increase in the permeability of the gut wall, which could allow bacteria or bacterial products such as the endotoxins (LPS) to enter the blood stream. The current model suggests that the

uptake of endotoxins into the blood stream induces systemic inflammatory diseases (Pollitt 2004; Levin et al. 1970).

Once the inflammation is induced in the horse, Hood et al. (1990) proposed that the pressure in the foot leads to ischemia and death of cellular structure in the connective tissue that attached the coffin bone to the hoof wall. Since the late 1990s, Christopher Pollitt, one of the leading laminitis researchers, has refuted this (Pollitt & Davies 1998). Pollitt argued that an increase in hoof temperature, which has been observed in laminitic horses, suggests that a greater amount of blood is reaching the hoof (Pollitt & Davies 1998). Therefore, the blood flow to the area would increase during the development and acute stages of laminitis, thus disproving the hypothesis that ischemia is initiates the development of laminitis (Pollitt & Davies 1998). When horses are exposed to a carbohydrate overload have their feet cooled in ice water, resulting in an 80.5% reduction of blood flow in the foot, laminitis does not develop (Pollit 2004; Pollitt and Davies 1998). Therefore, laminitis is likely not the result of ischemia. Rather, it is likely due to factors released into the blood stream by either hindgut bacteria or by the host immune response to the shifts in microflora induced by the increase in nonstructural carbohydrates (Pollitt 2004).

Current research suggests that the most likely causative agents to the laminae and basement membrane in the hoof are a group of enzymes called the gelatinases (Asplin et al 2007; Johnson et al. 1998; Mungall & Pollitt 1999; Mungall et al. 2001; Pollitt 2004; Kyaw-Tanner et al 2008). These endopeptidase enzymes are a part of a large protein family known as the matrix metalloproteinases, which are dependent on calcium and contain zinc (Gomis-Ruth et al. 1997). The general role of the MMP proteins are to degrade the extracellular matrix during normal tissue reorganization and repair (Shapiro 1998; Pollitt 2004). However, abnormal expression of MMPs are linked several different diseases in both humans and domestic animals, including

arthritis, tooth decay, cardiovascular disease, and tumor progression (Shapiro 1998). Gelatinases are responsible for the enzymatic degradation of collagen. Collagen is a component of the basement membrane (Barsky et al. 1983). It has been shown that laminitic horses have significantly higher levels of MMP2 and MMP9 when compared to healthy horses in blood serum and plasma during the development of laminitis (Mungall & Pollitt 1999). An increased number of MMP2 transcripts were also observed during the development phase of laminitis in mid-dorsal wall laminar tissue (Kyaw-Tanner & Pollitt 2008). A histology similar to laminitis has been induced *in situ* by subjecting cross-sections of hooves to APMA (p-aminophenylmercuric acetate), a known MMP activator (French & Pollitt 2004). Batimastat (BB-94), a known inhibitor of MMPs, is given to horses it prevents laminitis from developing (Pollitt et al. 1998). This may indicates that the overexpression of MMP in the hoof either causes the development of laminitis or at the very least play a significant role in its progression (Mungall & Pollitt 1999).

MMPs are produced by granulocytes during a normal immune response; however, over-activity causes damage (Lindberg et al. 2001; Newby 2007). Gelatinases, specifically MMP9, contribute to the destruction of the connective tissue, which leads to the separation of the coffin bone from the hoof wall (Pollitt 2004). The presence of these MMPs could be an immune response related to sepsis or endotoxemia. If laminitis where to be induced by endotoxemia or sepsis, the symptoms should be present throughout the animal (Pollitt 2004). However, disruption of the connection between the epidermis and the dermis is only observed in the hoof (Pollitt 2004). One possible explanation is that the connective tissue in the hoof is under considerably more pressure than in other parts of the body. This added stress could tear the

coffin bone from the hoof wall if the connective tissue has been compromised by sepsis induced MMP activity (Pollitt 2004).

One of the lead hypotheses is that the cascade that leads to the production of MMPs and laminitis is initiated by excessive consumption of NSCs as a result of the subsequent shift in the GI microflora, change in metabolism, and acidosis of the hindgut. The increase in the permeability of the intestinal wall allows for either bacteria or bacterial products to be absorbed into the blood stream. The subsequent immune response leads to the over expression of MMP and the degradation of the BM, causing laminitis (Pollitt 2004). To better understand the factors in the hindgut that contribute to this disease it is important to first characterize the healthy response of the hindgut microflora to an increase NSCs consumption.

Characterization of Bacterial Communities

Diversity Indexes

Many of the ecological indices that are currently used to characterize microbial communities were originally developed for use in traditional ecology, referring to the study of the ecology of macro-ecosystems. There are some known problems when these metrics are applied to microbial communities (Schloss & Westcot 2011). Bacterial taxonomy outlines are built on what is known from cultured species; however, the sequencing of 16S rDNA has led to the realization that the bacterial kingdom is far more diverse. This is due to the fact that many bacterial species cannot be cultured (Schloss & Westcot 2011; Lozupone & Knight 2005). Because many of these bacteria have never been cultured, many candidate taxa such as TM7 and taxa that are rarely able to be cultured such as *Acidobacteria* do not have complete taxonomy outlines (Schloss & Westcot 2011). Because of this it is very hard to systematically define criteria for many bacterial species as well as broader taxonomic classifications.

Because species level identifications are only possible for species that have been extensively cultured, most bacterial studies utilize a classification called an Operational Taxonomic Unit (OTU). This is often used as a synonym for species (Schloss & Westcot 2011).

Alpha diversity

Alpha diversity indices include OTU /species richness, the *chao1* index of bacterial richness, as well as the *Simpson's D* and the *Shannon's H* values (Koleff et al. 2003; Sepkoski 1988). Next-generation sequencing has enable a more comprehensive analysis of microbial diversity and is allowing researchers to observe rarer bacterial taxa that have never been observed before (Schloss & Westcot 2011). However, how the sequence data is used to define taxa can be an issue. There are two popular methods to identify bacteria. The first is to sort bacterial sequences based on similarity to a reference sequence of an already described species (Schloss & Westcot 2011). The second method sorts reads based on their similarity to each other. The groups that are formed from this method are known as operational taxonomic units or OTUs and share at least 97% sequence identity (Schloss & Westcot 2011; Daly et al. 2001). The OTU method addresses some of the issues that are present when reads are just mapped to reference sequences (Schloss & Westcot 2011). When reads are sorted by OTUs rather than phylotypes, it allows every read to be sorted based on the same criteria. When reads are binned by similarity to a reference, it restricts the researcher by only allowing the reads to be sorted into already observed taxas (Schloss & Westcot 2011). With the OTU method, regardless of whether or not the species had been described, it can still be placed into an OTU and data would not be lost. Once OTUs are formed, they can then be placed in a family or genus of bacteria utilizing reference sequences this is often referred to as phylotyping.

OTU richness counts the number of unique taxa observed. Another very similar index is *chao1*. In most natural populations the number of OTUs surveyed directly increases with the sampling effort (Hughes et al. 2001). An accurate OTU richness can be difficult to obtain in microbial studies. This is primarily due to sampling bias due to the fact that different bacteria are more readily sequenced than others (Hughes et al. 2001; Acinas et al. 2005; Kanagawa 2003). The c*hoa1* index applies a corrective factor based on the relative frequency of the OTUs observed to the number of OTUs that observed in a sample.

C*hao1* applies a capture-mark-recapture method to add a corrective factor based on sample effort to the OTU richness observed.

$$Chao1=S_{obs} + (n_1^2/n_2^2)$$

Where $n_{1\ is}$ observed once in the data set by the number of OTUs that were captured at least twice (n2). This value is squared and added to the number of observed OTUs (S_{obs}) (Chao 1984). This is a very useful tool for microbial studies because the index was designed for datasets that are skewed towards a large proportion of low observation OTUs (Hughes et al. 2001). *Chao1,* along with OTU richness, are simple indices that describe the diversity as a function of presence and absence, but do not provide information on the distribution of the diversity.

Shannon's H and Simpson's D both describe the relative evenness of OTU abundance. Both incorporate information on how many times an OTU was observed but they are based on two different approaches (DeJong 1975). The Shannon's H index assumes that all species in the population are present in the dataset and that sampling was random. Estimation of evenness therefore only pertains to the observed OTUs. This assumption may not hold in microbial communities because obtaining a random sample of bacterial communities can be difficult because of PCR bias for certain bacterial lineages (Hughes et al. 2001; Acinas et al. 2005;

Kanagawa 2003). Shannon's H handles low abundance OTUs better than Simpson's D. Simpson's D gives greater weight to higher abundance OTUs. This leads to low abundance OTUs being underrepresented (DeJong 1975). Microbial communities are often dominated by a few high abundance OTUs. As was the case in Manfield et al. (2002), each index handles this in a different way. For instance, Simpson's D will not be affected as much by OTUs that were observed at low frequency, and thus will be more representative of the dominant and more biologically significant species. Shannon's H will give a more representative estimate of evenness, but due to the inclusion of low abundance OTUs, it may skew the data by including bacteria that are not a substantial portion of the community.

Beta Diversity

Whereas alpha diversity is useful for describing the diversity within a population, beta diversity describes the diversity, or the differentiation, between communities. Beta diversity can be based on the presence or absences of species, such as with the Sørenson and Jaccard indices of group overlap and the LibShuff metrics (Baselma & Orge 2012; Magurran 2004; Lozupone and Knight 2005). Metrics like these tend to lose valuable information (Lozupone and Knight 2005). When these metrics are applied to studies where bacteria are identified by the mapping of reads of an OTU to reference sequences, common practice mandates that for a positive species-level identification the sequences must have at least a 97% similarity with that of a known reference sequence. Meaning, regardless of their phylogenetic similarity to the reference, they are all treated the same way. These metrics do not take into account the evolutionary distance between populations (Lozupone and Knight 2005). For example, if two fecal samples were dominated by different, but closely related *Streptococcus* species with a 96% similarity, the aforementioned metrics would show that these two communities were more distant from each

other than they really are. In order to correct for this, distance based metrics such as the Unique Fraction index have been developed.

The Unique Fraction distance (*UniFrac*) matrix, is an approach for creating a distance matrix from the branch lengths within a phylogenic tree, created with all present OTUs found in all of the communities being analyzed (Lozupone and Knight 2005). The algorithm then identifies unique fractions of the tree that are made up of OTUs from each specific sample. The distance between different samples is a function of the branch length between their unique fractions. If no unique fractions between communities are found, the samples are given a distance value of 0. If the samples are completely unique and do not share any of the same OTUs, they are given a distance value of 1 (Lozupone and Knight 2005). The significance of the distance matrix is tested with Monte Carlo simulation. Additionally, *UniFrac* values can either be weighted, where the relative abundance of the OTUs is taken into consideration, or unweighted, where the abundance is disregarded and singletons and high abundances OTUs are treated in the same way (Lozupone and Knight 2005). *UniFrac* distances can then be used in conjunction with multiple statistical tests, including but not limited to UPGMA and PCoA, to provide a more comprehensive comparison of bacterial communities (Lozupone and Knight 2005).

Justification of Research

Regardless of the exact pathology of the disease, each proposed carbohydrate induced route to laminitis starts with the over consumption of NSCs. It is assumed that this causes a shift in the hindgut microflora, however, this shift has not been closely characterized nor are the functional implications well understood. In recent years, newer sequencing methods have been developed, by companies such as Roche and Illumina that allows a vast amount of OTUs to be

sequenced and identified. The hindgut microflora of horses have been characterized many times (Al Jassim and Andrews 2009; Al Jassim 2011; Van den Berg et al. 2013; Daly et al. 2001; Dougal et al 2013). Though the gut microflora of horses has been extensively characterized, many of these studies have often focused on specific bacteria such as *Streptococcus* ssp. and *Lactobacillus* ssp. There have been relatively few studies applying next-generation sequencing to the study of the hindgut microflora in horses (Costa et al. 2012; Steelman et al. 2012; Dougal et al 2013). However, high throughput sequencing technologies are becoming more accessible and are starting to be applied more. The cost of these technologies is decreasing, specifically, the Illumina MiSeq platform provides a greater amount of reads for less cost than the older Roche 454 FLX. It has been shown that the Illumina MiSeq is more accurate than the Roche. Despite this, to date there has been only 1 published study applying Illumina sequencing to the identification of equine hindgut microflora (Moreau et al. 2014). This thesis will compare sequencing of horse fecal microflora by both a Roche 454 FLX and an Illumina MiSeq. This will be used to determine which sequencing platform is the best suited for further investigations of equine hindgut microflora.

The traditional method of determining the community makeup of the hindgut microflora is to either culture or sequence bacterial DNA obtained from the feces (Van den Berg et al 2013; Steelman et al. 2012; Mallett et al. 1983; Ley et al. 2008). This is primarily due to the ease of collection and because it does not affect or injure the animal in any way. Yet, the substantial microbial activity occurs in the cecum of the large intestine. To date, the difference in cecal and fecal communities has not been examined. Because the cecum is where the majority of the fermentation takes place, it should logically be the site that is most affected by the change in the pH due to carbohydrate overload. Additionally, as the lactic acid and VFAs leaves the cecum, it

is possible that it would be diluted by the contents already in the lower part of the large intestine as well. Because of this, there may be functional difference in the GI microflora of the cecum and the feces. Even though it may be convenient and less invasive to analyze the microflora of the feces, it may not be representative of the cecum. In order to determine whether the fecal samples are appropriate for analysis of hindgut microflora, fecal and cecal samples collected at the same time will be compared utilizing a Roche 454.

In conclusion, there will be two major outcomes from this thesis. First, to determine whether the Illumina MiSeq protocol or the Roche 454 FLX protocol is better suited for the study of the equine gut microflora. Second, whether feces or cecal contents provide a more representative sample of the hindgut microflora. Together these conclusions will be used to determine an appropriate method for further surveys of the microflora of the equine hindgut.

Chapter 2. Material and Methods

Study animals and Experimental Treatments

Seven American quarter horses with cecal cannulas were used for the feed trial performed by Dr. Josie Coverdale, AUP Permit #2011-24 (Texas A&M University Institutional Animal Care and Use Committee). Horses were castrated males weighing between 497-580 Kg. Individuals were paired by weight and randomly distributed into two experimental groups in a crossover design. One group was given a low carbohydrate diet of commercial feed (LC, 0.6% of body weight) and the other a high carbohydrate diet (1.2% of body weight). The commercial feed used was Vitality Perform 14 Horse Feed (Cargill Animal Nutrition, Elk River, MN) (table 1). The total daily amount given to each horse was separated evenly into two meals administered at 6:30 and 18:30. Between treatments, the horses were allowed to consume coastal Bermuda grass hay (*Cynodon dactylon*) *ad libitum* (Table 1). At the end of a 28-day period, each experimental group was fed costal Bermuda grass for a 28-day wash out and then the alternate treatment was given to each group for another the 28 days.

TABLE 1. Relative concentrations of nutritional components in the feed used during this study. The total nonstructural carbohydrates consist of both sugar (specific type not identified) and dietary starch. Nonstructural carbohydrates make up a maximum of 38% of this feed. In addition, the maximum percentage of Neutral detergent fiber (NDF) and Acid Detergent Fiber (ADF) are included in this table as well.

Item	Feed	Hay
Dry Matter , %	89.9	91.9
Acid Detergent Fiber %	5.9	36.2
Neutral Detergent Fiber %	14.3	70.1
Crude Protein %	14	9.6
Non Structural Carbohydrate %	30	
Calcium, %	0.9	0.4
Phosphorous, %	0.6	0.2

Samples

Cecal samples were collected via a cecal cannulation. Fecal samples were collected immediately after a bowel movement. The collection occurred at 0h (the 6:30 am feeding was used to designate 0h), 3h, 6h, 9h and 12h on Day 1, at 0h, 6h and 12h on Day 2 and 3, and at 6 h on the remaining days of the treatment. 7.30 ml of cecal contents or feces were collected. Samples were placed on ice and within an hour of collection the samples were frozen (-20^{o} C) until they were processed. Due to the costs associated with sequencing only a subset of the total 252 samples were analyzed. The following samples were used to compare Illumina and 454 methods and cecal versus fecal sampling: Day 1 at 0 h (n = 14), Day 3 at 6 h (n = 14), and Day 7 at 6 h (n = 14).

Isolation of Microbial DNA

Microbial DNA from fecal and cecal samples was extracted at Department of Veterinary Integrative Biosciences, Texas A&M University (Warzecha 2013) using the QIAamp DNA Mini Stool Kits (Qaigen, Valencia, CA) with modification to the manufacturer's protocol. Samples were thawed while vortexing. Once the samples thawed, they were allowed to vortex for another 5 minutes in order to suspend the material. Approximately 3 ml of the material was placed into 15 ml conical tubes to which 3ml of ASL buffer was added. The mixture was centrifuged at 500 x g for 1 minute and 500 ul was transferred into a 2 ml screw cap tube with 0.15 mm garnet beads (Mo Bio Laboratories Inc., Carlsbad, California) and 1 ml of phenol:chloroform:isoamyl alcohol (25:24:1) . The tubes were vortexed at 1,500 rpm for 5 min, incubated at 95^{o}C for 5 minutes, and centrifuged at 13,000 x g for 1 minute. The top aqueous 0.4 ml of the supernatant was transferred into a new 2 ml snap cap tube with a 1 ml of ASL buffer. An InhibitEx tablet was added and the tubes were vortexed at 1,600 rpm for 5 minutes. The suspension was incubated at room temperature for 1 minute and centrifuged at 16,000 xg for 7 minutes. All of the supernatant was transferred into a new 1.5 ml tube, centrifuged again, and supernatant was re-transferred to a 2 ml tube with 25 ul of ProteinaseK and 540 ul of buffer AL lysis buffer. This was briefly vortexed, centrifuged and incubated at 70^{o}C for 20 minutes.

After incubation, 540 ul of 100% ethanol was added, tubes briefly vortexed, and centrifuged and 540 ul of the ethanol/lysate solution was added to QIAamp column, centrifuged at 16,000 xg for 1 minute. The lysate was discarded and columns were placed in a new collection tube and steps repeated. After all lysate passed through column, 500 ul of wash buffer AW1 was added and tubes were centrifuged at 16,000 xg for 1 minute. 500 ul of wash buffer AW2 was added to each column, and centrifuged at 16,000 xg for 4 minute. Spin columns were placed in a new

collection tube centrifuged again at 16,000 xg for 2 minutes to ensure all of the AW2 buffer had been removed. The columns were placed in a sterile 1.5 ml snap cap tube and 200 ul of elution buffer was added. The tubes were allowed to incubate at room temperature for 5 minutes. The columns were centrifuged at 16,000 xg for 1.5 minutes and the elutant with the DNA was transferred into a 0.5ml screw cap tube and properly labeled. DNA was stored at -20˚C. In order to determine the quality of the DNA was electrophoresed on a 1% agarose gel stained with ethidium bromide and quantity estimated with a NanoDrop. All samples were diluted to 20 ng/µl in water. Samples with less than 20 ng/µl DNA were concentrated using ethanol precipitation.

Amplification of DNA

A 600-bp segment of the 16s rRNA gene was PCR amplified. Cecal and fecal samples were amplified with the universal Eubacterial primer 530F (5'-GTGCCAGCMGCNGCGG-3') and 1100R (5'-GGGTTN CGNTCGTTG-3') (Tay et al. 2002), barcoded, and sequenced on the Roche 454 FLX (454 Life Sciences). These primers target the v4 and the v6 region of the 16s rRNA gene. To compare the 454 method with the more recent Illumina protocol, the fecal samples were also analyzed using the Illumina 16S diversity assay. For this procedure, the same fecal DNA were amplified with the S–D-bact primer set; S-D-Bact 0341F (5-CCTACGGGNGGCWGCAG-3) and S-D-Bact 0785 R (5-GACTACHVGGGTATCT AATCC-3) (Klindworth et al 2012), following the Illumina protocol (San Diego, California) in the 16s Metagenomics Sequencing Library Preparation Manuel (Part # 15044223 Rev. B). These primers target the V3 and V4 regions of the 16s DNA yielding a 460-bp amplicon. The amplicons were purified using Agencourt Ampure Xp beads (Beckman-Coulter; Brea California) and indexed with the Nextera Indexing Kit, and again bead purified. Samples were normalized and pooled per manufacturer recommendations and sequenced Illumina MiSeq (2 x 300 bp

reads) at the Genomics Core Facility at Cornell University (Ithaca, New York). A total of 40 cecal and 40 fecal samples from matching collection points were sequenced with the 454. The 40 fecal samples were also sequenced using the Illumina protocol. Samples were standardized and pooled. All reads analyzed were single reads, not paired ends. Reads were phylotyped against the GreenGenes database v 13.5 (updated January of 2016 http://greengenes.lbl.gov).

Data Analysis

The microbial diversity observed in the samples were compared between the two sample types (cecal versus fecal) and sequencing methods (454 versus Illumina). Reads with mean quality score < Q30 were removed and 20 bp off the 3' end Illumina sequences were also trimmed in CLC Genomics Workbench v 8.1 (Qaigen Bioinformatics, Hilden, Germany). Sequences were exported in .fna format and imported into Quantitative Insight into Microbial Ecology (QIIME) v 1.9.1 (Caporaso et al., 2010). The .fna files were combined and reads from each sample demultiplexed using the script below.

```
add_qiime_labels.py -i fasta_dir -m example_mapping.txt -c InputFileName -o
combinedseqs.fna
```

The output seq.fna file was used for the subsequent analysis. The OTUs were identified using a 97% similarity criteria with the uclust algorithm and singletons were excluded. Representative OTU sequences were compared to the GreenGenes database v 13.5 and a taxonomy assignment was made (Price et al. 2009; Edgar 2010). The script for these functions is below.

```
pick_de_novo_otus.py -i seqs.fna -o otus
```

To compare the alpha diversity among the samples the *chao1*, OTU richness, species richness, *Simpson's D*, and *Shannon's H* were calculated with the below script.

```
alpha_diversity.py -i otu_table.biom -m PD_whole_tree, chao1, chao1_ci,
otus_observed, species_obsreve, simpson, shannon -o adiv_pd.txt -t
rep_set.tre
```

The *Simpson's D* and *Shannon's H* were used to evaluate relative evenness within the microbial communities. OTU richness, species richness and *chao1* provide information on the total number of distinct taxonomic groups including those that are yet to be assigned to a taxa. In order to compare differences in number of reads per samples, rarefactions were run using 100, 1,000, 5,000, and 10,000 randomly selected reads based on both *chao1* and OTUs and rarefaction curves were plotted using the script below

```
alpha_rarefaction.py -i otu_table.biom -o arare_max100/ -t rep_set.tre -m
Fasting_Map.txt -e 100
```

Beta diversities were estimated between all samples. Beta diversity values were averaged within sample sets (Roche fecal, Roche cecal, and MiSeq fecal) and between sample sets. This was done using the weighted UniFrac method incorporating the phylogenetic tree constructed when identifying the OTUs. This was done with the following script.

```
beta_diversity.py -i otu_table.biom -m weighted_unifrac -o beta_div/ -t
rep_set.tre
```

The mean UniFrac estimates for each sample set (Roche fecal, Roche cecal, and MiSeq fecal) were used to determine how diverse they were. In addition, weighted UniFrac distance was used to compare fecal vs cecal samples and the two sequencing methods. Principle Coordinate of Analysis (PCoA) was used to visualize the divergence between samples and groups. The script used is below.

```
beta_diversity_through_plots.py -I otu_table.biom -o bdiv_even100/ -t
rep_set.tre -m Fasting_Map.txt -e 100
```

The changes that occurred in the microflora in response to treatments were compared between the MiSeq and 454 data sets as well as the fecal and cecal samples identified using the

Roche 454 FLX sequencer. A paired Student's t-test was used to test for significant differences for mean number of OTUs identified, *chao1*, Simpson's D, Shannon's H, weighted UniFrac values, and taxa summaries.

Chapter 3. Comparison of Two Sequencing Strategies; the Roche 454 FLX and the Illumina MiSeq

Comparison of the reads produced by each platform

In order to determine the efficacy of each sequencing platform used in this study, the number of usable reads that each platform produced was analyzed. It is important to note that the only difference between the two sets of samples being compared in this section are the protocols used to sequence them. Otherwise, the samples are the same and DNA extractions are identical. This allows the two sequencing methods to be compared with each other. This comparison is made only for sequences of high enough quality to be assigned to an OTU. Overall, the Roche 454 produced more consistent results, although it produced far less reads than the Illumina MiSeq (Table 9 in appendix).

The MiSeq produced a mean of 17,987.1 reads sample (range: 739 to 142,894, S.D. = 9,248.8) (Table 2). The two extreme outliers (mj100 and mj111) were excluded from the above estimates. In contrast, the 454 had a mean of 8,638 reads per sample (range: 2,154 reads to 36,349, S. D. = 3,943). Although the 454 had a smaller standard deviation than the Miseq it also had significantly less reads than the MiSeq ($P=7.51X10^{-7}$). The standard deviations were normalized as percentages of the mean for a more direct comparison. The standard deviation of the 454 was 32.6% of the mean; whereas the standard deviation of the MiSeq was 52.4%. There were no extreme outliers in the 454 data set.

TABLE 2. Comparison of summary statistics for the 454 and MiSeq methods including the minimum and maximum number of reads observed in a sample, the median and mean number of reads per sample, and the standard deviation, the mean OTU richness, *chao1*, Shannon's D, the Simpson's H, standardized Simpson's D. All listed P-values estimated using the Students' T-tests. The *chao1* upper and lower bounds represent a 95% CI.

MiSeq		454		Significance (T-Test)
Min:	739	Min:	3237	
Max:	142894	Max:	14678	
Median:	16777	Median:	9049	
Mean:	17987.08	Mean:	8779.03	7.51E-07
Std. dev.	9424.83	Std. dev.	2863.35	
OTUs	3449.95	OTUs	2430.05	3.24E-04
Simpson's D	0.99	Simpson's D	0.99	0.45
Shannon's H	9.25	Shannon's H	9.37	0.39
Stand. Sim.'s D	2.87E-4	Stand. Sim.'s D	4.1E-4	0.47
chao1	11014.29	chao1	5373.41	7.95E-08
chao1 lower bound	10038.48	chao1 lower bound	4974.31	
chao1 upper bound	12153.54	chao1 upper bound	5837.77	

Comparison of alpha diversity between the two sequencing

OTU and Species Richness

The MiSeq detected a significantly greater OTUs richness than the 454 (Student's T-test, $P = 3.24X10^{-4}$) with 3,449.95 OTUs per sample compared to 2,430.05 for the 454 (Table 2, Figure 2). Therefore, the MiSeq was able to identify more distinct taxonomic groups in the same set of fecal sample than the 454. To determine if the number of OTUs was a function of the additional reads produced by the MiSeq, the number of reads was plotted against the number of OTUs per sample for both methods (Figure 3). When samples with equal number of reads are

compared, the 454 detects more OTUs; however, the MiSeq is capable of generating more reads and therefore can detect more OTUs. The 454 has a more even distribution of reads and OTUs per sample.

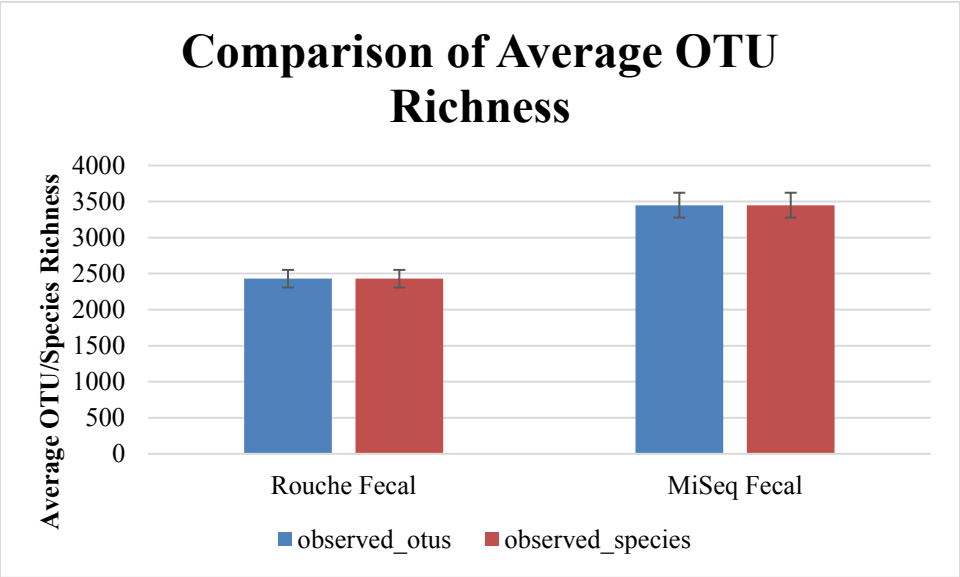

FIGURE 2. The mean OTUs per sample for the MiSeq or the Roche 454 for 40 fecal samples. The error bars represent a 95% CI.

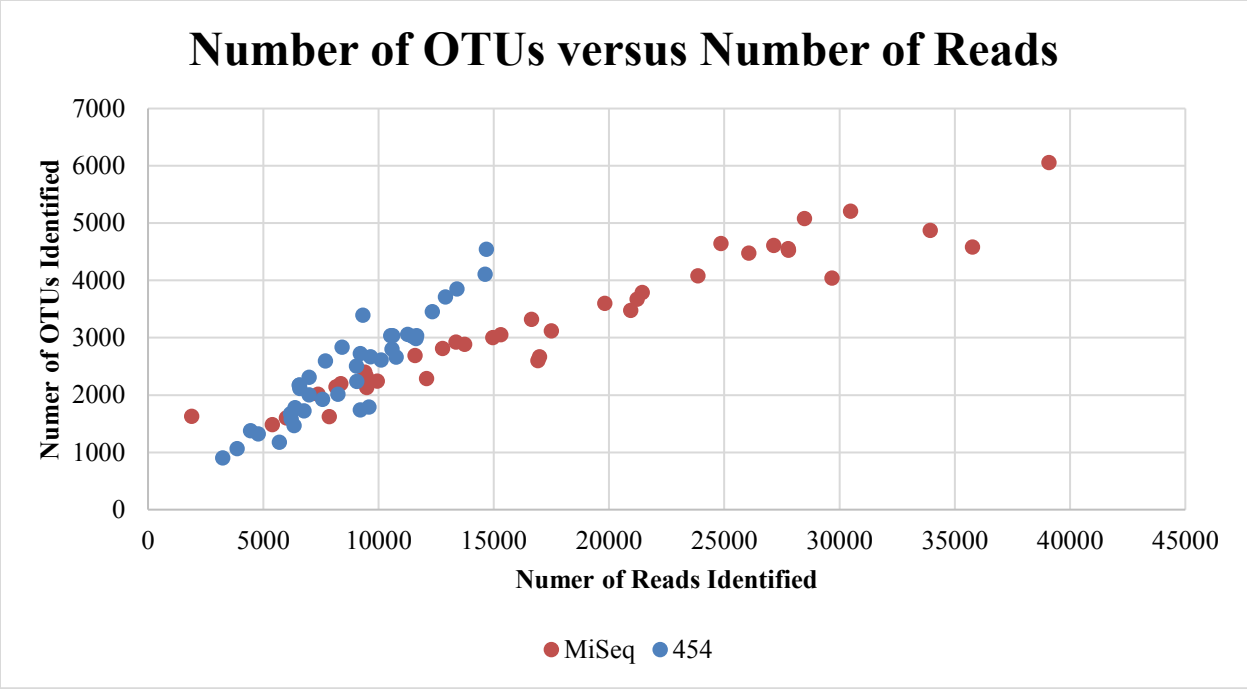

FIGURE 3. The number of OTUs per sample plotted against the number of reads.

Comparison of *Chao1* Values

The *chao1* values were directly proportional to OTU richness observed in the data sets. The mean *chao1* value for the 454 fecal data was 5,373.42 (95% C.I. = 4,974.3-5,837.85; whereas the mean for the MiSeq was 11,014.29 (95% C.I. = 10,038.5 – 12,153.4; Table 2, Figure 4). The confidence intervals do not overlap, indicating that the mean MiSeq *chao1* is significantly greater, which is confirmed by a two-tailed paired Student's T-test (P-Value=$7.95X10^{-8}$).

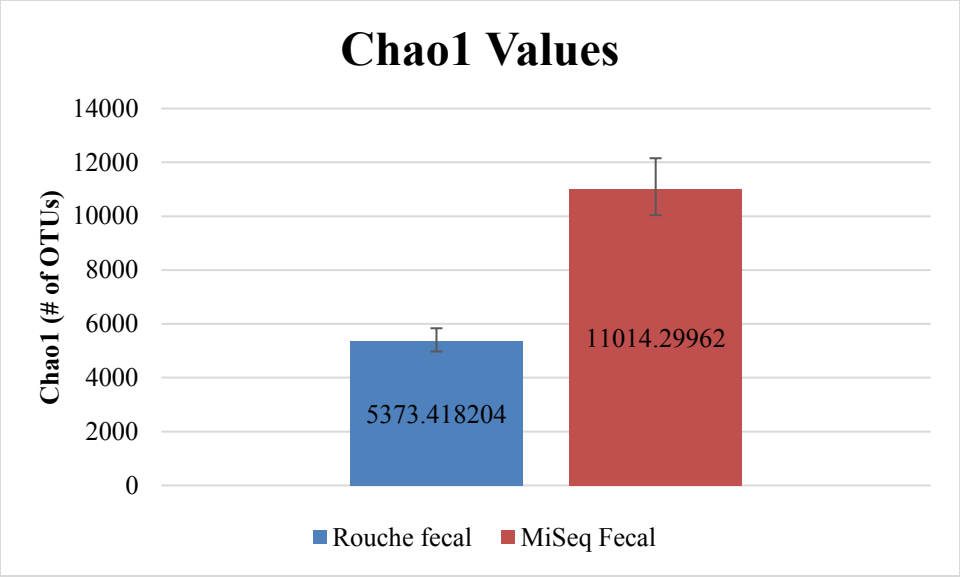

FIGURE 4. C*hao1* values for the MiSeq and the 454 fecal samples. The error bars represent 95% confidence intervals.

Pairwise Comparison

Four fecal samples exhibited a similar number of reads when sequenced using the MiSeq and the 454 sequencing protocols. Sample 4 exhibited 9,479 reads in the MiSeq data set and 9,203 reads in the 454 data set. Sample 5 exhibited 15,201 reads in the MiSeq data set and 14,678 reads in the 454 data set. Sample 54 exhibited 11,586 reads in the MiSeq data set and 12,904 reads in the 454 data set. Sample 63 exhibited 8,364 reads in the MiSeq data set and 9,051 reads in the 454 data set. Reads identified in the MiSeq and the 454 data set were not significantly different (P=0.58). The *chao1* and the OTU richness were not significantly different between the two data sets (P=0.11; P=0.44) (Table 3).

TABLE 3. A pairwise comparison of fecal samples with a similar number reads identified between the MiSeq and the 454 data sets. Relevant metrics such as the number of reads identified, the OTU richness and the *chao1* values are included in this table.

Pairwise Comparison of Fecal Samples						
Sample Number	Number of Reads Identified		Number of OTUs		*Chao1*	
	MiSeq	454	MiSeq Data	454 Data	MiSeq Data	454 Data
4	9479	9203	2322	2725	7855.91	5443.67
5	15301	14678	3055	4545	8692.02	9112.38
54	11586	12904	2688	3712	7672.09	8282.1
63	8364	9051	2200	2236	5695.6	4580.7
Significance	0.58		0.11		0.44	

Alpha Rarefactions

In order to determine the efficacy of each platform to identify all OTUs found in the samples, alpha rarefactions were conducted for both OTU richness and *chao1* with 1,000 and 10,000 randomly selected reads for each MiSeq sample. The rarefaction curve for 1000 reads begins to show a decreasing rate in the discovery of OTUs/read around 1000 reads, although it still has not plateaued (Figure 5a). The sample MJ6 is uncharacteristic because of its steep slope at 1000 reads. The rarefaction curves for *chao1* are similar to the OTU curves (Figure 5b). Similar patterns were observed for the MiSeq 10,000 read rarefactions, The OTU richness is still steadily increasing and while the curve is beginning to plateau, it has not yet reached a plateau. (Figure 6a and 6b).

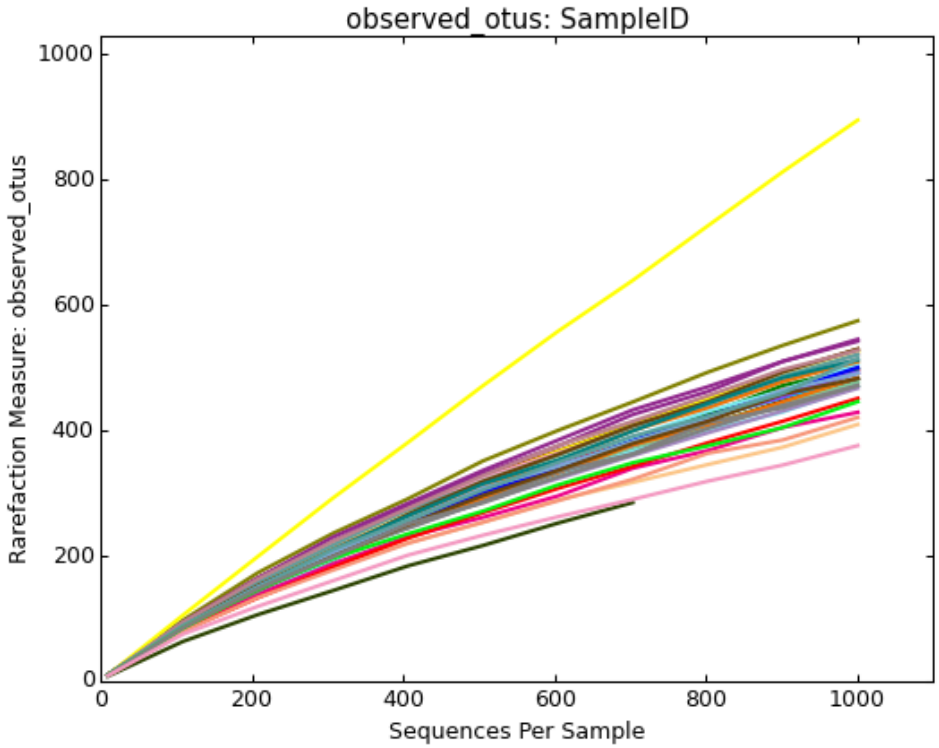

FIGURE 5A. Rarefaction curve produced by randomly sampling 1,000 reads from each sample in the MiSeq data set.

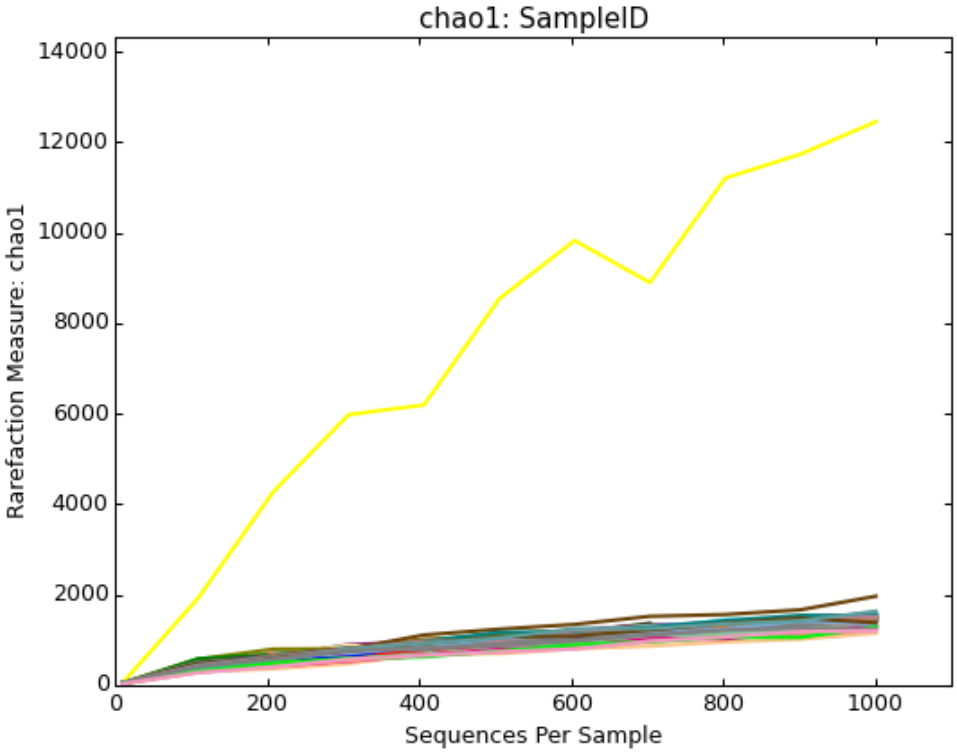

FIGURE 5B. *chao1* rarefaction curve produced by randomly sampling 1,000 reads from each sample in the MiSeq data set.

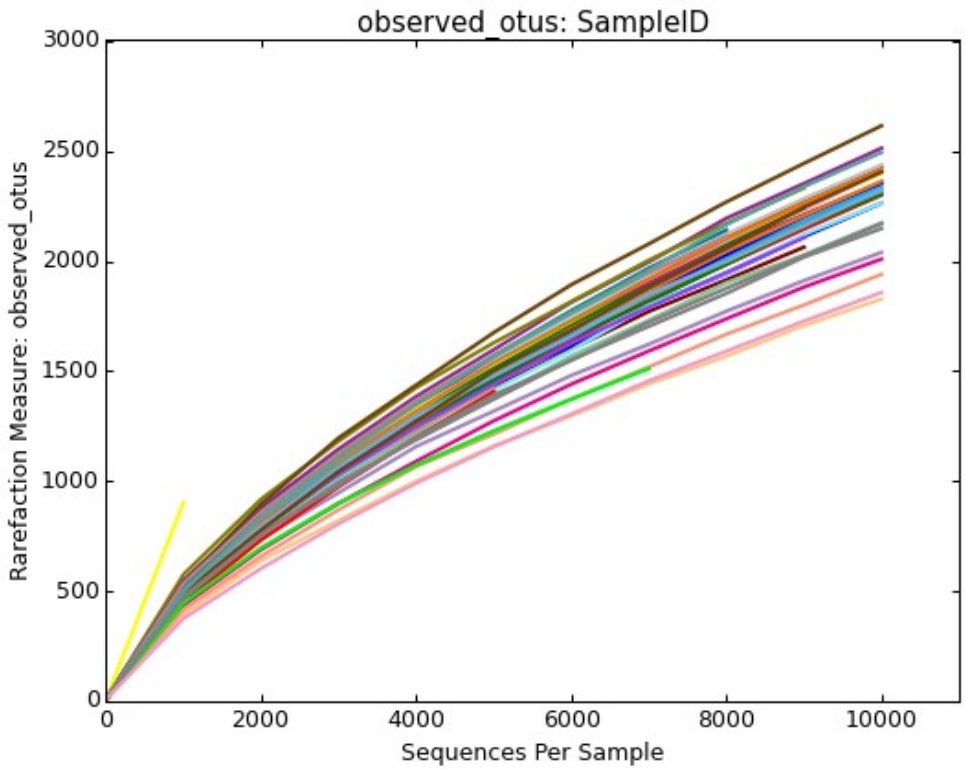

FIGURE 6A. OTU rarefaction curves for 10,000 randomly selected MiSeq reads.

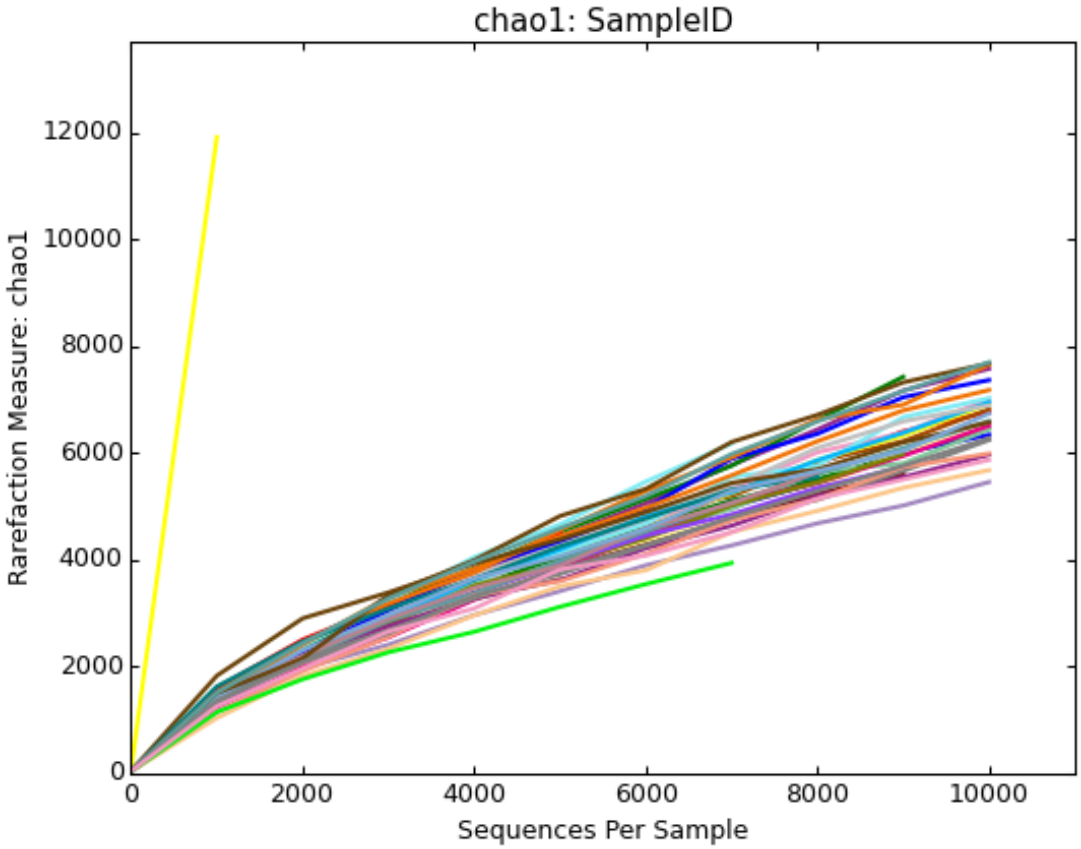

FIGURE 6B. Rarefaction curve of the *chao1* value for 10,000 randomly selected MiSeq reads per sample.

The 454 fecal samples exhibited similar patterns when rarified at 10,000 random reads for the OTU richness to the MiSeq data set, but for the *chao1* value of the 454 appeared to plateau after around 1,000 reads (Figure 7). The majority of the 454 samples have *chao1* values less than 6,000, whereas in the MiSeq data approximately half of the samples were able to reach a *chao1* value of at least 6,000 OTUs.

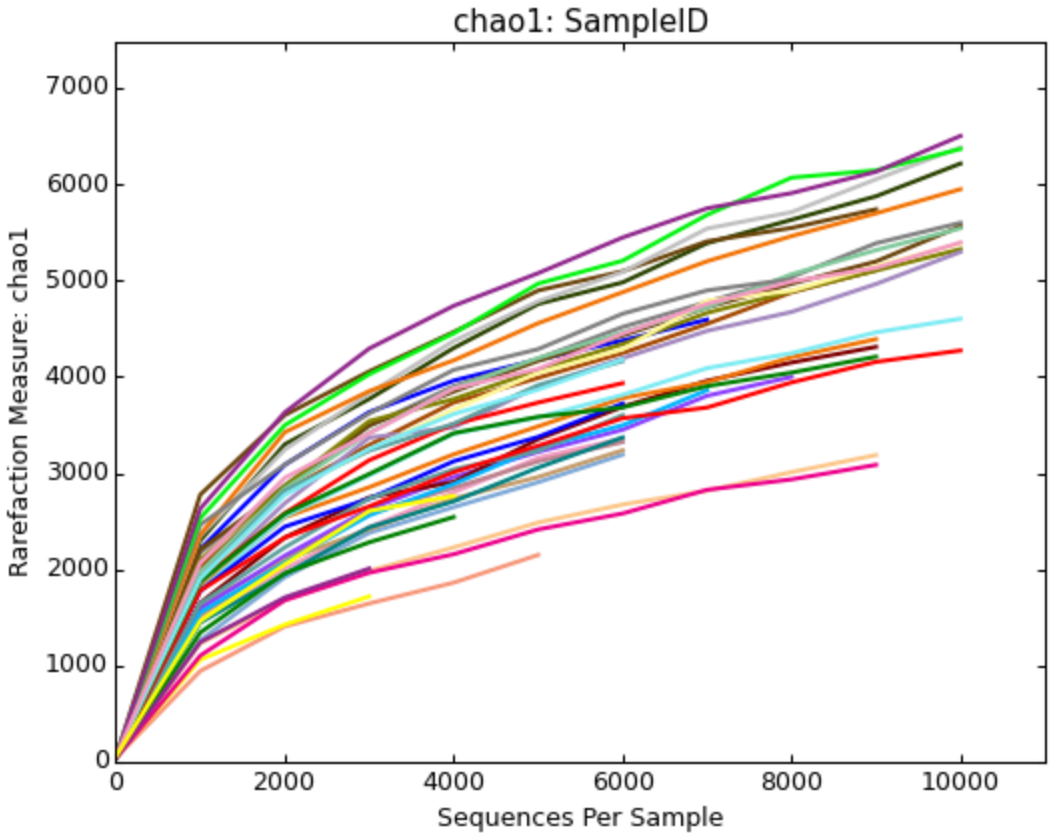

FIGURE 7. *chao1* rarefaction for 10,000 randomly selected reads using the 454 data.

Comparison of the Relative Evenness of the OTUs Identified

Simpson's D and Shannon's H values were used to compare the evenness of the sampled microbial communities by comparing the relative abundance of the OTUs. This section focuses on Simpson's D because it is easier to standardize and compare among communities. Simpson's D can range from virtually 0, indicating near complete dominance by 1 taxa, to the total number of species observed in the sample, indicating complete evenness in the community. By dividing the Simpson's D with the corresponding OTU richness for each sample, the values can be scaled so they are more directly comparable. The average Simpson's D value identified in the 454 data set was 0.093. In the MiSeq was 0.091 and the standardized Simpson's D (SSD) were 4.09X10-4 and 2.87X10-4, respectively (Table 2 and Figure 8). The Simpson's D and the SSD values

were not significantly different based on a two tail paired Student's T-Test (P = 0.45 and 0.47, respectively). Similar results were obtained with the Shannon's H, with an estimates for the 454 and MiSeq of 9.37 and 9.25 respectively (Table 2) (P = 0.39). Therefore, the two methods detected OTUs at similar relative abundances yielding comparable estimates of evenness. Both estimators did indicate uneven communities.

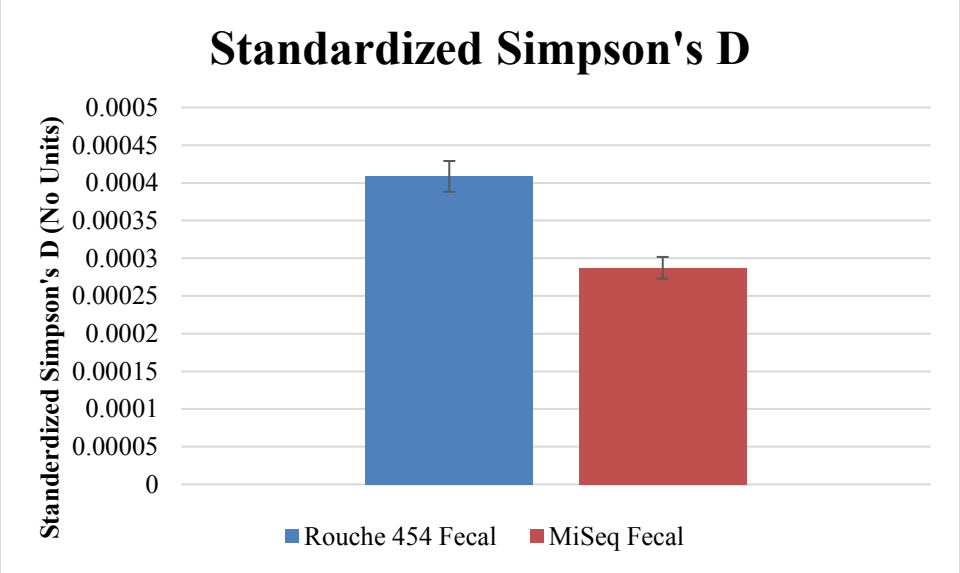

FIGURE 8. Standardized Simpson's D values for all samples estimated using the 454 and MiSeq methods. (P=0.45).

Comparison of the Beta Diversity

Weighted *UniFrac* was used to examine beta diversity. This provided information on the level of divergence between the OTUs observed in the MiSeq fecal and the Roche 454 fecal data sets (table 10 in appendix). The mean weighted *UniFrac* among the 454 samples was estimated to be 0.363 (SD =0.149) (Table 11 in appendix). The mean MiSeq *UniFrac* estimate was 0.349 (SD = 0.164). The mean *UniFrac* value between the MiSeq and Roche 454 data sets was 0.552 (SD = 0.104). This indicates that there is a methodological effect; the community diversity

estimated from the MiSeq and 454 fecal data show greater divergence than the variation among the different data sets. This is revealed when the, weighted *UniFrac* values analyzed and plotted with Principle Coordinate of Analysis (PCoA) (Figure 9).

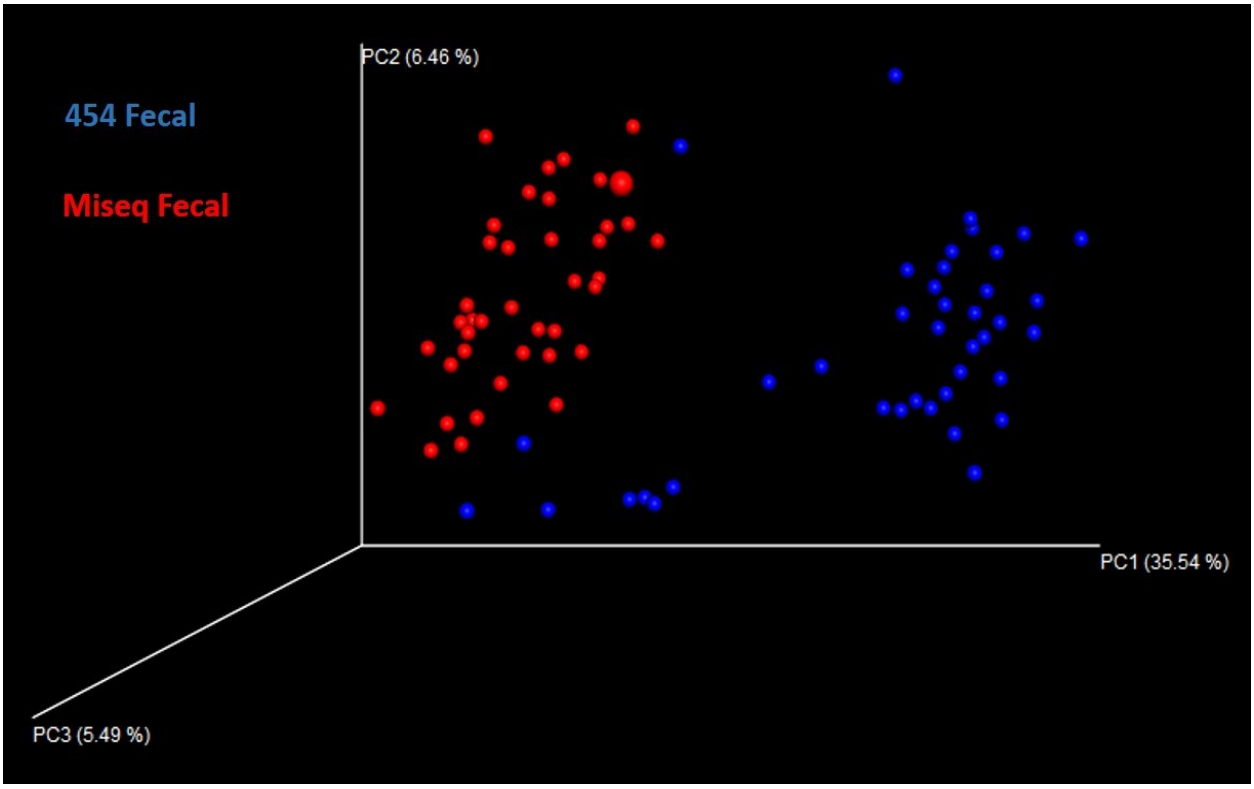

FIGURE 9. PCoA of the weighted UniFrac values.

The PCoA confirmed that the MiSeq and the Roche 454 fecal samples were substantially different from one another. The mean weighted *UniFrac* value suggests approximately 55% of the observed OTUs are unique to either the MiSeq or the 454. The PCoA identifies three principle coordinates, which are three statistical factors that explain the majority of the variation among the *UniFrac* estimates. The first principle coordinate explains 35.54 % of the total variation and this component is what separates the 454 and MiSeq data in the PCoA plot. Therefore, most of the variation in beta diversity is explained by the method used to assay the

diversity. Unweighted UniFrac values were also calculated and show similar trends to those observed in the weighted UniFrac values (Table 11 in appendix)

Comparison of Core Microbiomes

QIIME was used to determine the core OTUs observed among samples from 454 and MiSeq. The core microbiome is defined by Shade and Handleson (2012) as being comprised of the members common to two or more microbial assemblages associated with a habitat (Turnbaugh et al., 2007; Hamady and Knight, 2009). By identifying the species that are consistently present in samples, one can better understand the communities in an environment. The core biome defined in this study was one that included species identified in 95% of the samples. Using this criteria, there were 14 core OTUs observed in the 454 data set and104 core OTUs in the MiSeq data set (Tables 4 and 5 and Figures 10 and 11; Tables 17 and 18 in appendix). Species level identifications were not obtained.

TABLE 4. Core OTUs identified in 95% of the Roche 454 fecal samples.

454 Fecal Core Biom			
Phylum	Family	Genus	Count
Bacteroidetes	unknown		2
	Paraprevotellaceae		2
	Prevotellaceae	Prevotella	5
Firmicutes	Ruminococcaceae		1
	Veillonellaceae	Phascolarctobacterium	4
Total			14

TABLE 5. Core OTUs identified in 95% of the MiSeq Fecal samples are listed in this table.

MiSeq Fecal Samples Core Biom			
Phylum	Family	Genus	
Bacteroidetes	unknown	unknown	14
	Paraprevotellaceae	Prevotella	1
		CF231	2
		YRC22	1
	Prevotellaceae	Prevotella	2
	RF16	unknown	1
	BS11	unknown	1
Cyanobacteria	unknown	unknown	2
Fibrobacteres	Fibrobacteraceae	Fibrobacter	1
Firmicutes	unknown	unknown	9
	Mogibacteriaceae	unknown	4
	Clostridiaceae	unknown	1
	Lachnospiraceae	unknown	10
		Pseudobutyrivibrio	1
	Ruminococcaceae	unknown	38
		Ruminococcus	3
	Veillonellaceae	Phascolarctobacterium	1
Proteobacteria	unknown	unknown	1
	Campylobacteraceae	Campylobacter	1
Spirochaetes	Spirochaetaceae	Treponema	4
TM7	F16	unknown	2
Verrucomicrobia	RFP12	unknown	4
Total			104

Two phyla, *Firmicutes* and *Bacteroidetes,* comprised most of the core biome. *Firmicutes* accounted for 67 of the 104 OTUs and *Bacteroidetes* for 22 OTUs. *Firmicutes* are a well-studied group of Gram-positive bacteria that include *Clostridia*, *Bacilli*, and *Mollicutes*. All *Firmicutes* were in the Class *Clostridia*. Forty-one of them belonged to the family *Ruminococcaceae*. In addition, there were 11 OTUs in the *Lachnospiraceae*, 4 in the *Mogibacteriaceae* family, and 1 in the *Veillonellaceae* family. There were 10 unidentified OTUs. In contrast the 454 data, had only 5 OTUs in the Phyla *Firmicutes* and only one was in the *Ruminococcaceae* family, with the other 4 belonging to the *Veillonellaceae* family (Table 4 and Figure11). This suggests the MiSeq protocol was able to identify a greater range of taxa in the *Ruminococcaceae* and the rest of the *Firmicutes*, when compared to the Roche 454 platform. One notable exception *Phascolarctobacterium* for which 454 identified 4 OTUs whereas the MiSeq only 1.

A substantial number of *Bacteroidetes* were also identified in both sample sets. These are a large group of Gram-negative bacteria that are commonly found in the intestines of animals. The MiSeq identified 22 as opposed to the 454, which was only able to identify 9. For the MiSeq, of those identified to the family level, 4 were identified as belonging to the *Paraprevotellaceae* family, 1 was identified as belonging to the family *BS11*, 2 were identified as belonging to the family *Prevotellaceae* and 1 was identified as belonging to the family *RF16*. Conversely, the 454 data set produced 5 OTUs belonging to the family *Prevotellaceae* and 2 OTUs belonging to the family *Paraprevotellaceae*; with the remaining 2 OTUs being unidentified.

The Roche 454 was able to identify the *Bacteroidetes* to the family level more consistently, but the Roche 454 was not able to identify as many OTUs as the MiSeq. The

MiSeq also identified more Phyla than the Roche 454 protocol. This may be a result of either primer bias during the amplification of the 16S rRNA gene or differences in the number of reads between the 2 platforms.

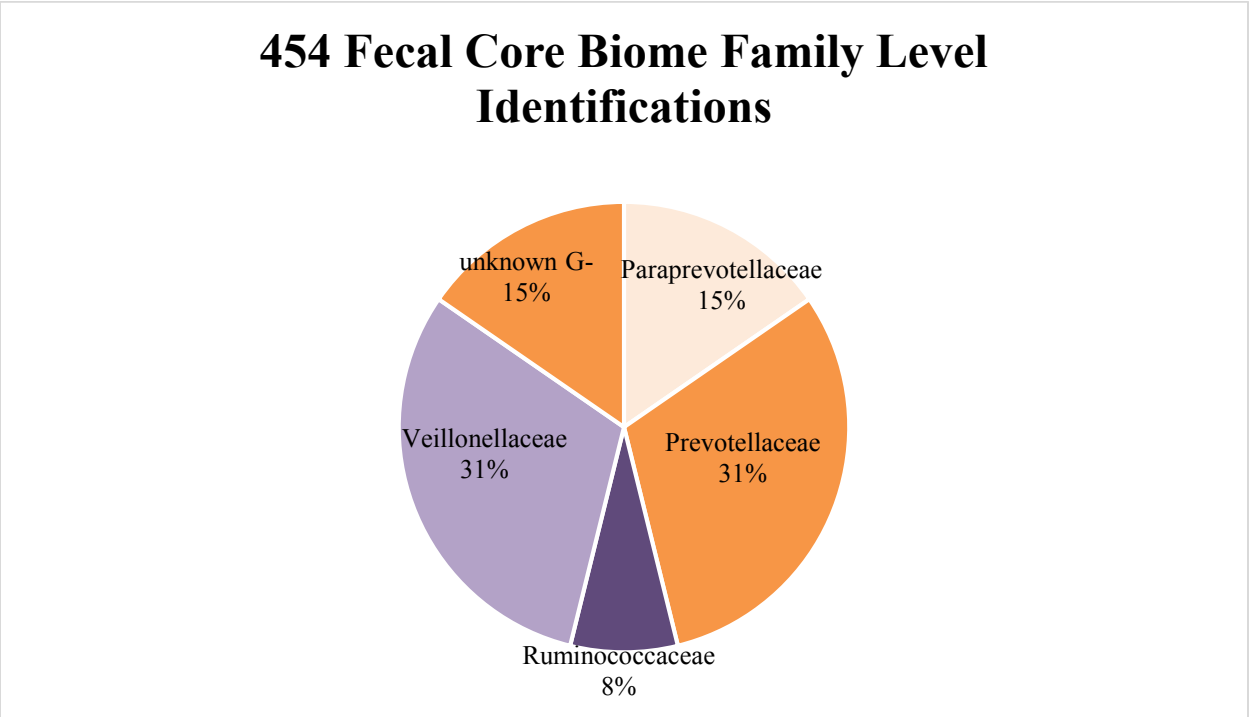

FIGURE 10. Bacterial families identified in the 454 fecal core biome.

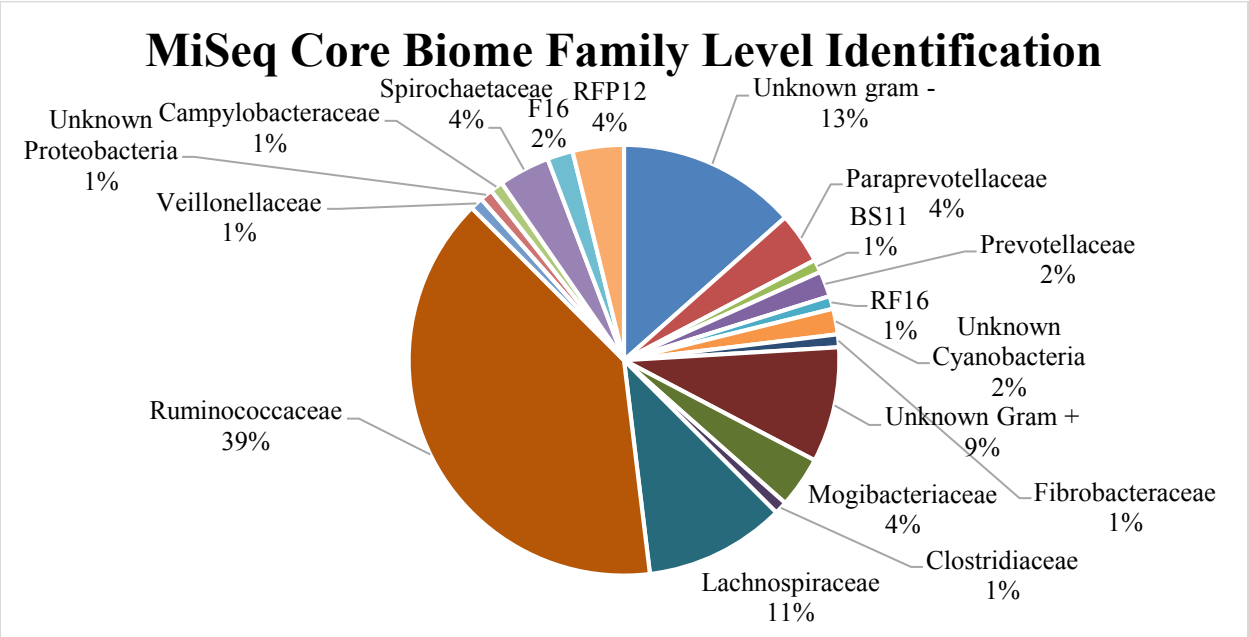

FIGURE 11. Bacterial families identified in the MiSeq samples.

Identification of Shifts in GI Flora by the Illumina MiSeq and the Roche 454 FLX

In order to conduct this analysis, QIIME was used to summarize the taxa present in each data set. Taxa were summarized as a percentage of the total amount of reads that were identified. Samples corresponding to different data sets, different treatments, and different days were separated. The percentage of reads that each taxa represented was averaged for each day; 1 (0h), day 3 (12:00), and day 7 (12:00) and for each treatment and each data set. The OTUs were sorted by percentage in decreasing order. The 10 most common OTUs were chosen for analysis. On average, the 10 most commonly sequenced OTUs accounted for 84.89% of the total reads identified.

The MiSeq data set did reveal some shifts in the GI tract of horse under the high carbohydrate treatment (Figure 12 and Table 12 in appendix). The first noticeable shift was in the total number of unassigned reads from 25.10% on day 1 to 5.28% on day 7. This decrease in unknown reads is accompanied by an increase in the amount of reads mapped to *Ruminococcaceae*, from 13.15 % to 20.27%. ($P = 0.18$). The abundance of *Paraprevotellaceae* reads also increased from 3.77% to 5.33% on day 7 ($P = 0.81$).

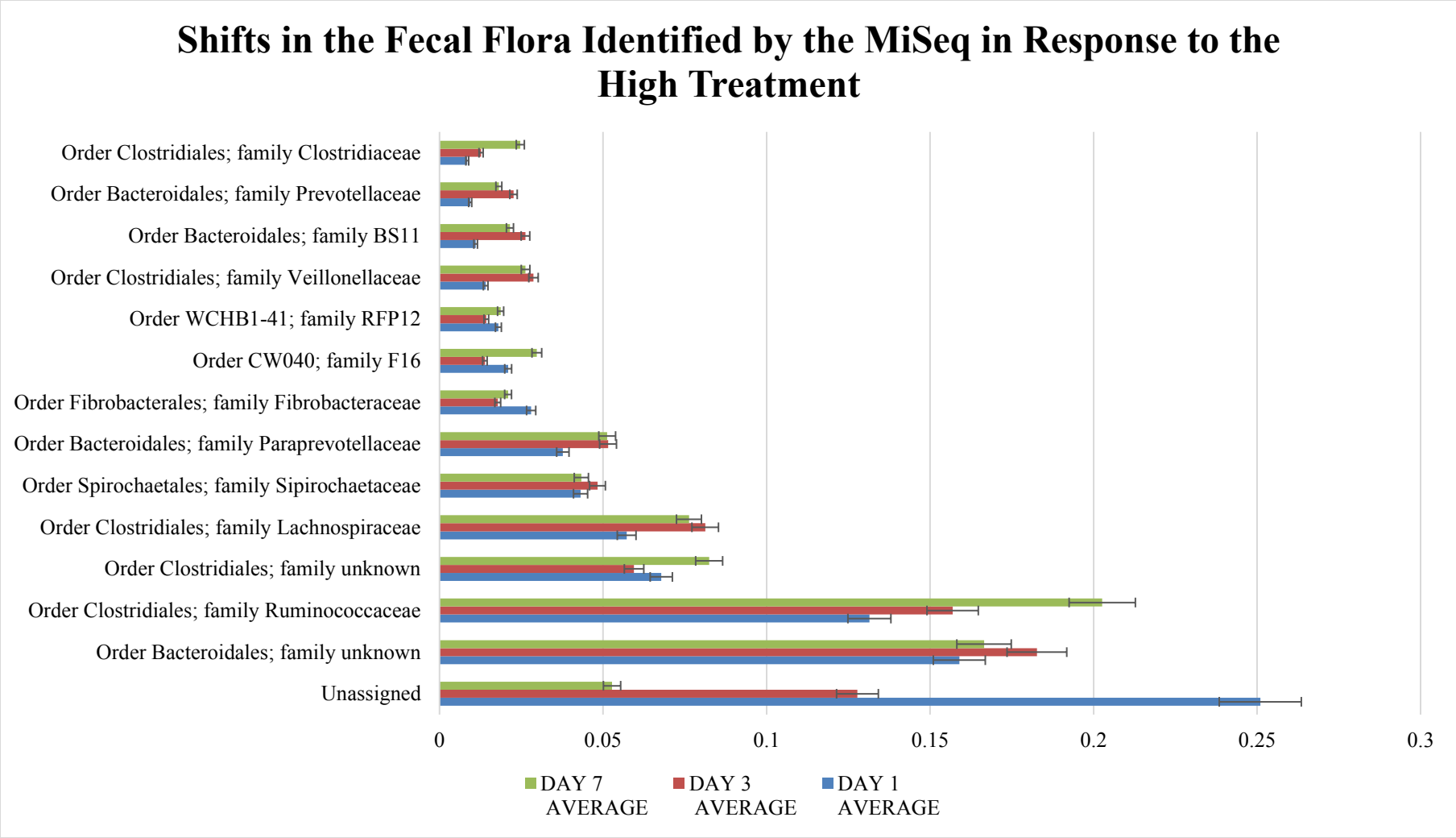

FIGURE 12. Shifts in the most prominent species of the GI microflora in response to 7 days on a high carbohydrate diet. The error bars represent 95% CI.

The 454 results were contradictory to the MiSeq data (Figure 13 and Table 13 in appendix). The relative abundance of unidentified reads did not change substantially. The 454 showed a decrease in *Ruminococcaceae* in response to the high carbohydrate treatment from 8.35% on day 1 and to 4.06% on day 7 (P=0.02). Increases in both *Paraprevotellaceae* and *Prevotellaceae* were also identified. Though *Paraprevotellaceae* did increase over the course of the treatment in the MiSeq data, the increase in the Roche 454 data is greater. *Paraprevotellaceae* accounted for 10.37% of the reads on the first day of the treatment and increased to 14.81% of the total reads identified on day 7 (P = 0.12). For *Prevotellaceae* there was an increase from 7.13% on the first day of the treatment and to 14.12% on day 7 of the treatment. This was not a statistically significant increase (P = 0.09). Finally, a minor increase in *Lachnospiraceae* OTUs was also observed, from 7.16% on day 1 to 9.95% on day 7 (P = 0.54)

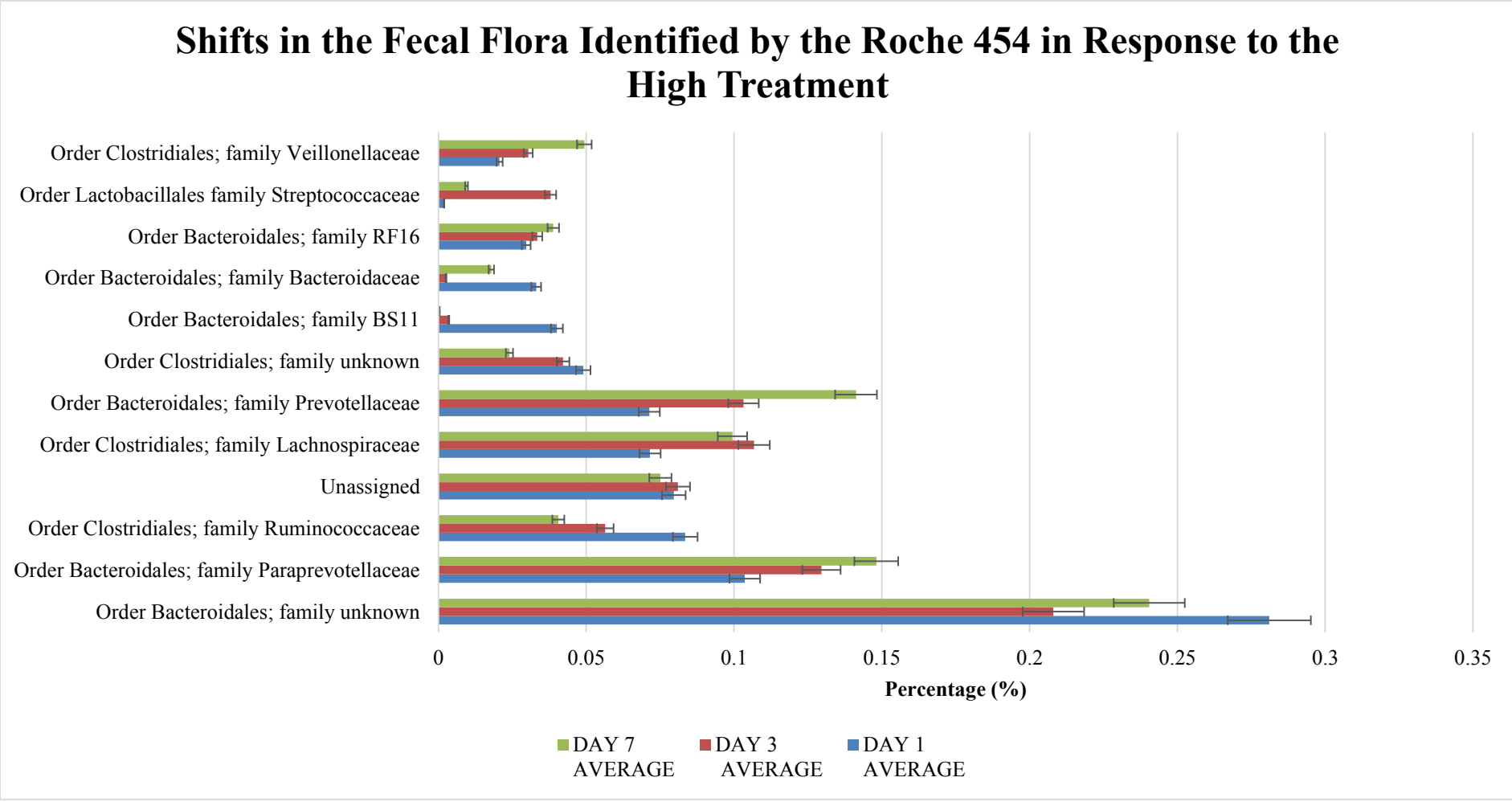

FIGURE 13. The shift in GI microflora in response to the high carbohydrate treatment using 454 data. The error bars represent 95% CI.

Chapter 4. Comparison of Microflora in Fecal and Cecal Samples using the Roche 454 FLX High Throughput Sequencer

Reads generated in each sample set

The mean number of reads for the fecal and cecal samples was not significantly different (P = 0.96), with 8,496.8 (S.D. = 1,959.4) and 8,779.0 (S.D. = 4,822.7) respectively per sample (Table 6 and Table 14 in appendix).

TABLE 6. Minimum and maximum number of reads observed in a sample, the median and mean number of reads per sample, OTU richness, and *chao1*, Simpson's D, and Shannon's H. for cecal and fecal samples analyzed using the 454. The lower and upper bounds of the chao1 value represent a 95% CI.

Cecal		Fecal		Significance (T-Test)
Min:	2154	Min:	3237	
Max:	36349	Max:	14678	
Median:	7798.5	Median:	9049	
Mean:	8496.8	Mean:	8779.03	0.96
Std. dev.	4822.72	Std. dev.	2863.35	
OTUs	2322.98	OTUs	2430.05	0.57
Simpson's D	0.99	Simpson's D	0.99	0.34
Shannon's H	9.17	Shannon's H	9.37	0.08
Stand. Sim.'s D	4.3E-04	Stand. Sim.'s D	4.1E-04	0.90
chao1	6769.29	chao1	5373.41	2.8E-03
chao1 lower bound	6145.25	chao1 lower bound	4974.31	
chao1 upper bound	7497.54	chao1 upper bound	5837.77	

Comparison of alpha diversity between fecal versus cecal

OTU and Species Richness

The mean OTUs in the fecal and cecal samples was similar with no significant differences (2,430.05 OTUs per sample and 2,322.98 OTUs per sample, respectively, P = 0.57) (Table 6 and Figure 14). The fecal and cecal samples had a similar maximum and minimum range in OTUs. This data suggests that the OTU richness is similar for fecal or cecal samples.

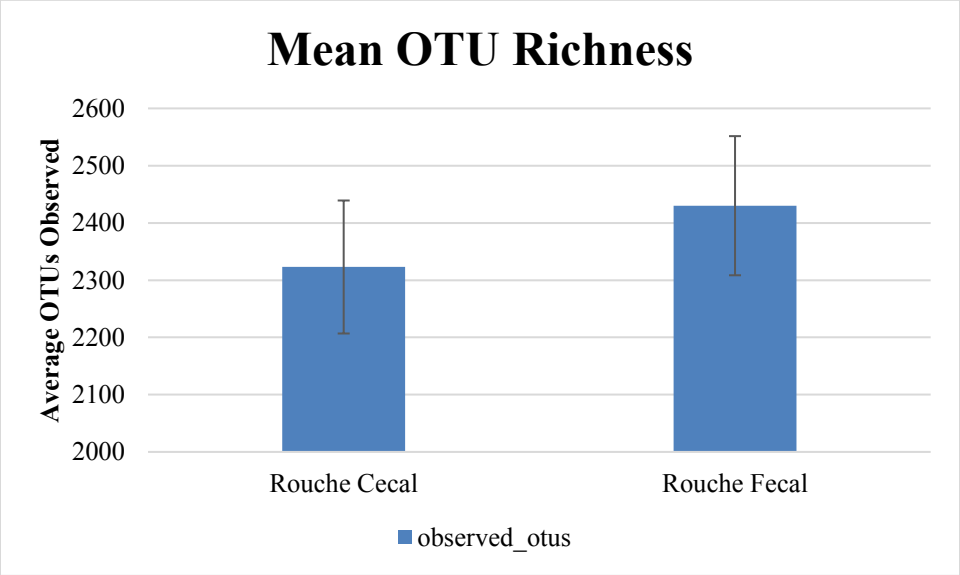

FIGURE 14. Mean OTU richness in the fecal or cecal samples. The error bars represent a confidence interval of 95%.

Comparison of Chao1 Values

Even though there was no significant difference in OTU richness, the *chao1* estimate was higher for cecal samples (6,769.29, S.D. = 2,187.01) compared to fecal (5,373.42, S.D. =1,861.59) (Figure 15). The difference in the *chao1* values was significant (Student's T-test, P = $2.8X10^{-3)}$.

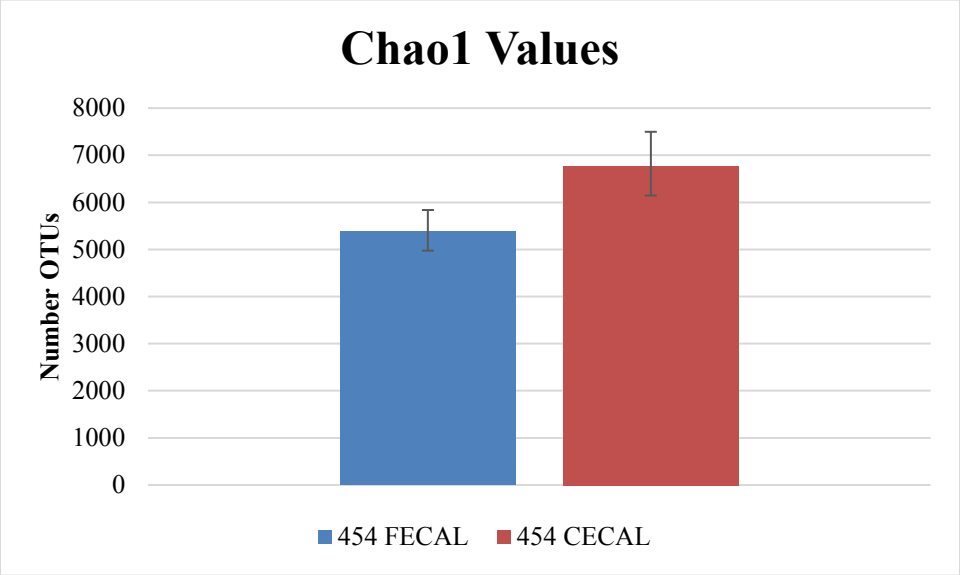

FIGURE 15. The mean *chao1* values of both the fecal and cecal datasets.

Alpha Rarefactions

Alpha rarefactions were conducted using 1,000, and 5,000, and 10,000 randomly selected reads; however only the 10,000 read rarefactions will be shown in this section. The maximum tested was 10,000 because this was just above the mean reads obtained per samples by the Roche 454 FLX sequencer (8,779.03 and 8,496.8 reads respectively).

At 10,000 reads the curves for the fecal and the cecal datasets have the same overall trend. Both exhibited a more substantial reduction in slope than was observed in curves built with fewer numbers of curves. The fecal dataset produced more consistent curves than those in the cecal dataset (Figure 16A and 16B; Figure 17A and 17B). The *chao1* curves exhibit a greater decrease in slope than the number of OTUs. This indicates observed OTUs in the samples were limited by the number of reads.

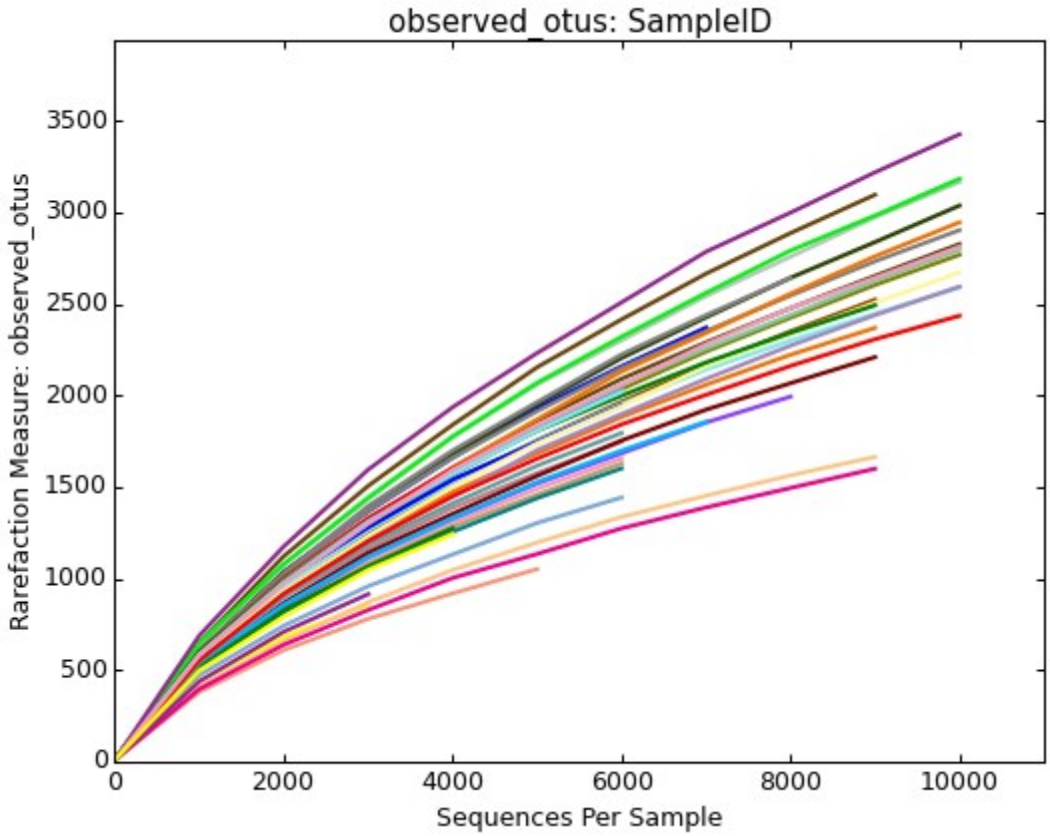

FIGURE 16A. Rarefaction curve for OTUs produced by randomly sampling 10,000 reads from each sample in the fecal dataset.

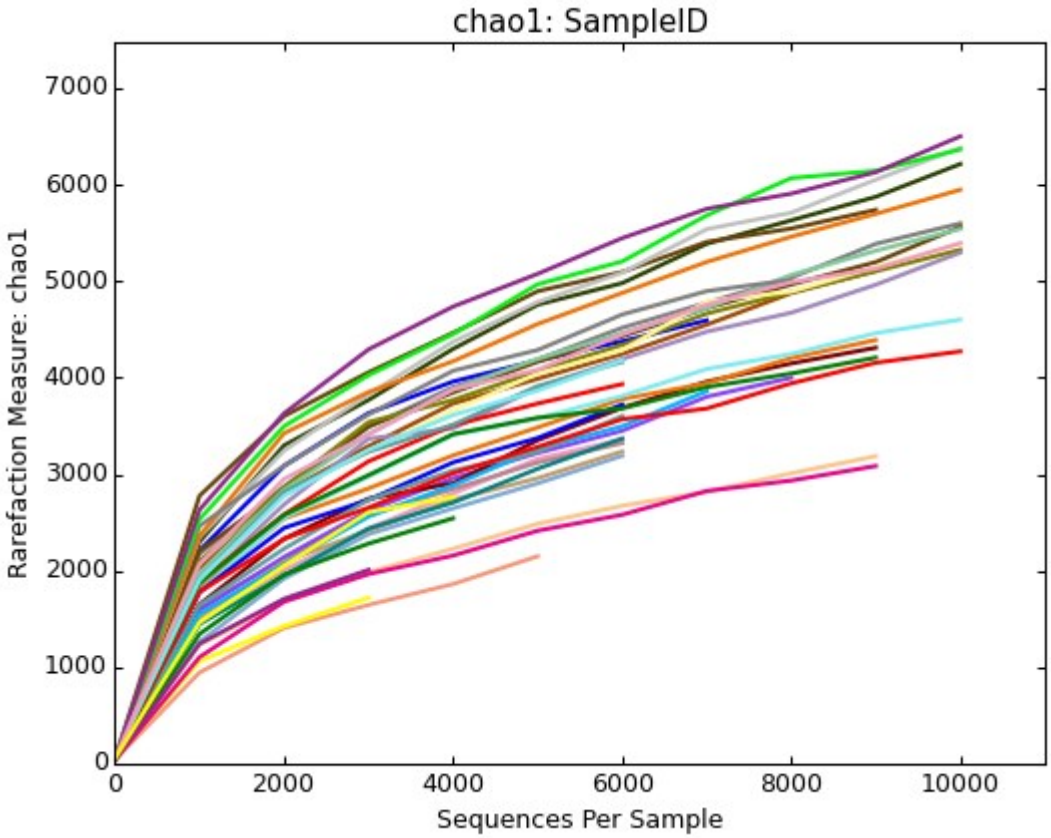

FIGURE 16B. Rarefaction curve for *chao1* produced by randomly sampling 10,000 reads from each sample in the fecal dataset.

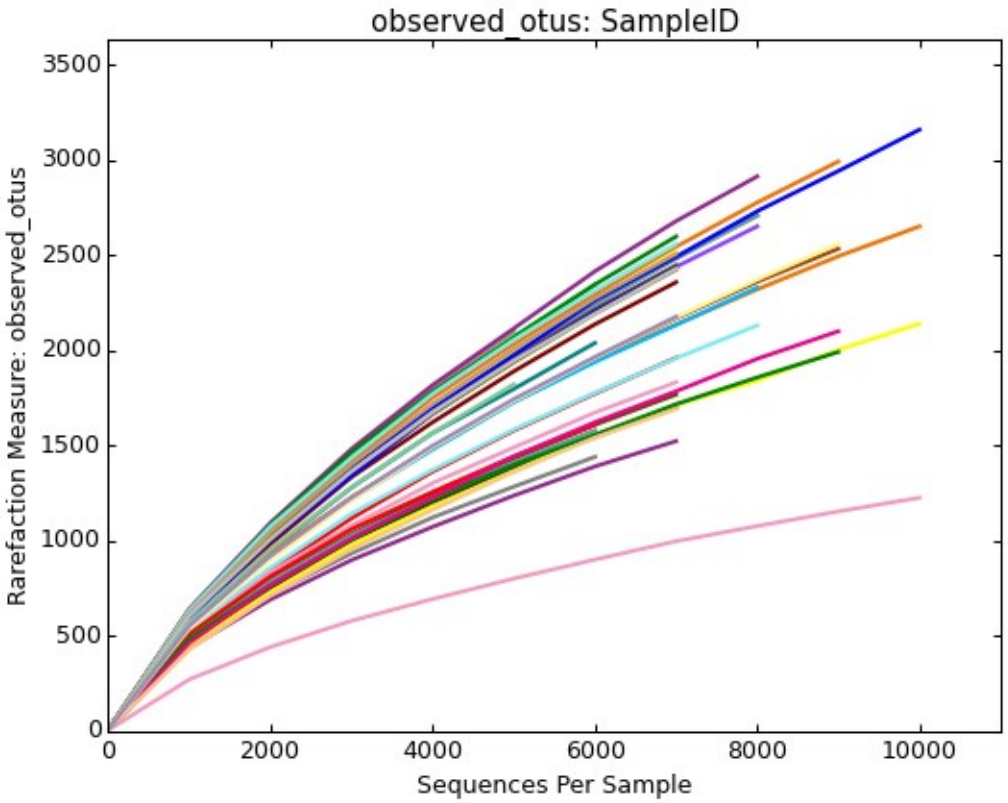

FIGURE 17A. Rarefaction curve of OTUs produced by randomly sampling 10,000 reads from each sample in the cecal dataset.

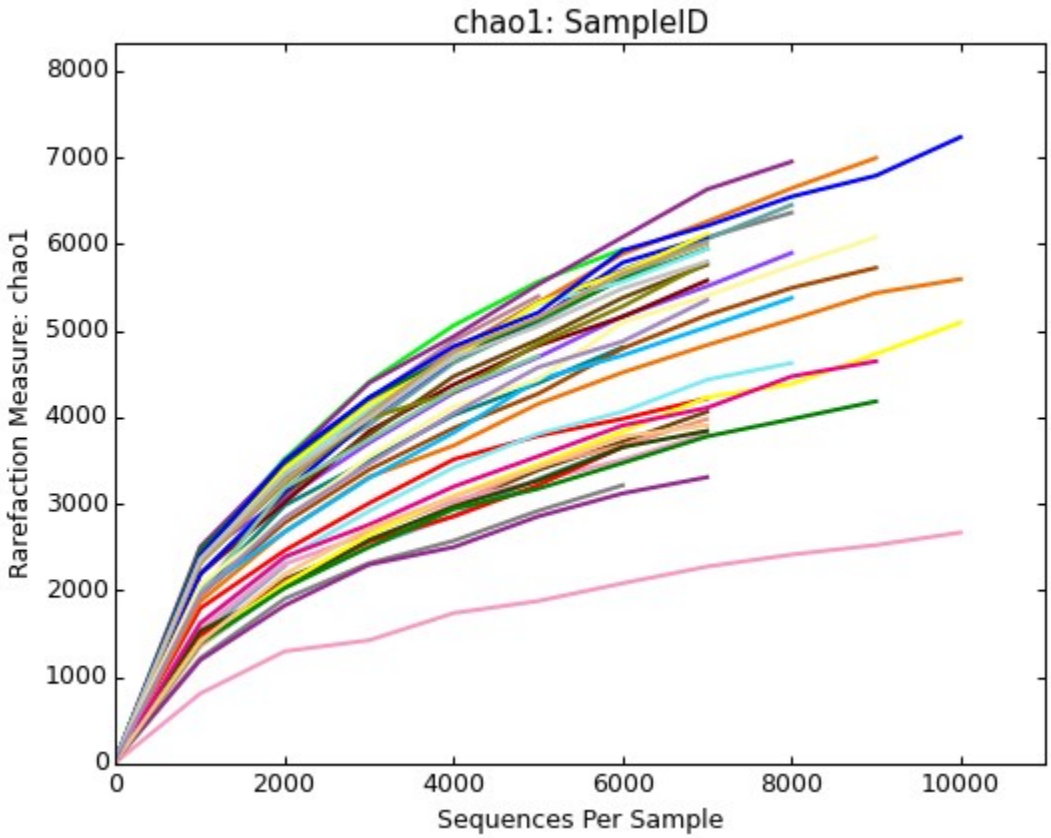

FIGURE 17B. Rarefaction curve of *chao1* produced by randomly sampling 10,000 reads from each sample in the cecal dataset.

Comparison of the Relative Evenness of the OTUs Identified

Simpson's D and Shannon's H values were used to determine the relative evenness of bacterial communities in the samples. A significant difference in the Simpson's D or Shannon's H would indicate one of them has an environment more conducive for specific bacterial species thereby enabling them to dominate the community.

The average Simpson's D value observed in the cecal dataset was 0.989 out of a possible score of 2322.98, which was the average OTU richness observed in this particular dataset (Table 9). The Simpson's D values were standardized (SSD) by dividing by the corresponding OTU

richness observed in the sample. The mean SSD for the cecal dataset was $4.26X10^{-4}$. For the fecal samples the mean Simpson's D = was $9.99X10^{-1}$, and SSD was $4.11X10^{-4}$. The Shannon's H values were comparable to the Simpson's D estimates. The fecal samples exhibited a Shannon's H of 9.37; whereas, the cecal samples were observed to have a Shannon's H value of 9.17. In both the Simpson's D and Shannon's H metrics, both cecal and fecal samples had similar estimates (P = 0.3.41 and 0.816). Therefore, uneven communities dominated by few OTUs that comprised the majority of their respective biomes were observed in both fecal and cecal samples.

Comparison of the Beta Diversity

The beta diversity between the fecal and the cecal samples was compared with weighted *UniFrac*. The mean weighted *UniFrac* value between fecal samples was 0.36 (S.D.=0.15) and in cecal samples 0.28 (S.D.= 0.11) (Table 10). This indicate that the fecal samples show more diversity among the OTUs relative to the cecal samples, although this difference is not significant because the overlapping standard deviations. Similarly, the mean *UniFrac* distance between fecal and cecal samples was 0.33 (S.D. = 0.14). This suggests that the beta diversity among the fecal and cecal samples was similar to that between them.

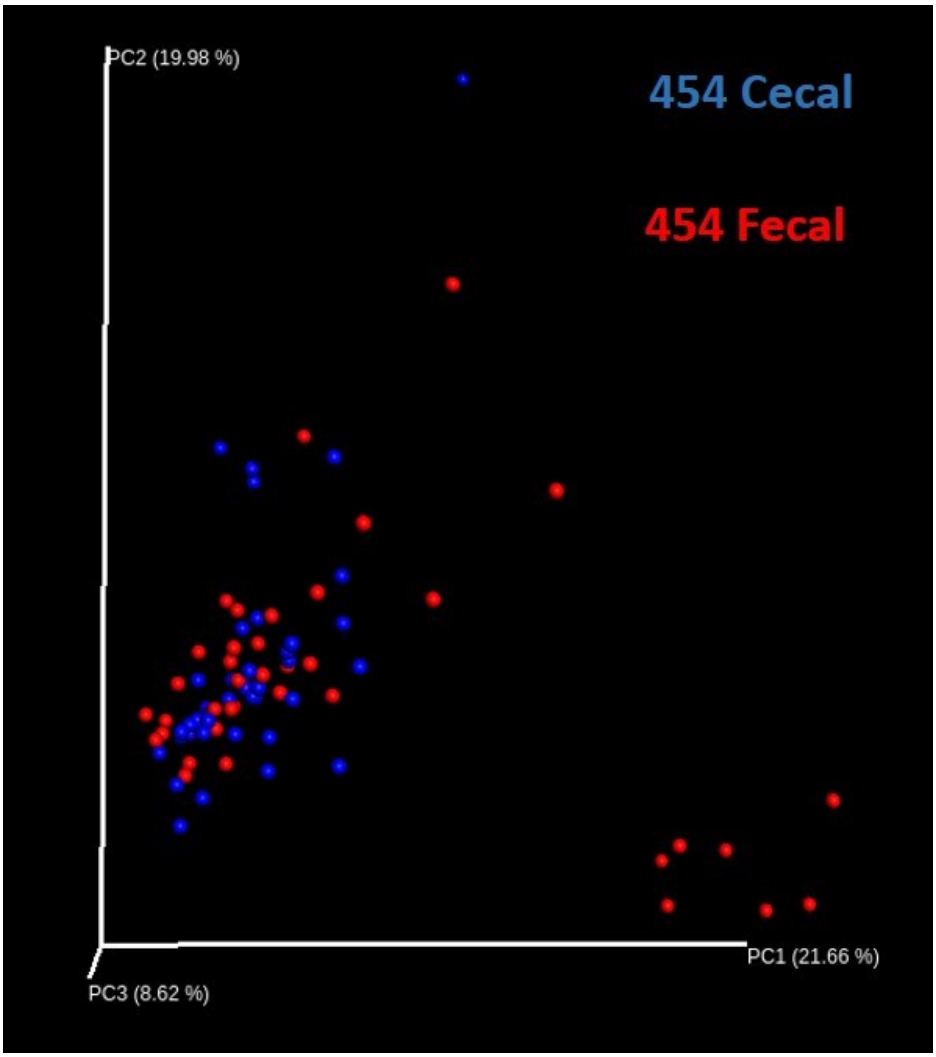

FIGURE 18A. PCoA of the weighted UniFrac values of all samples used in this experimentation.

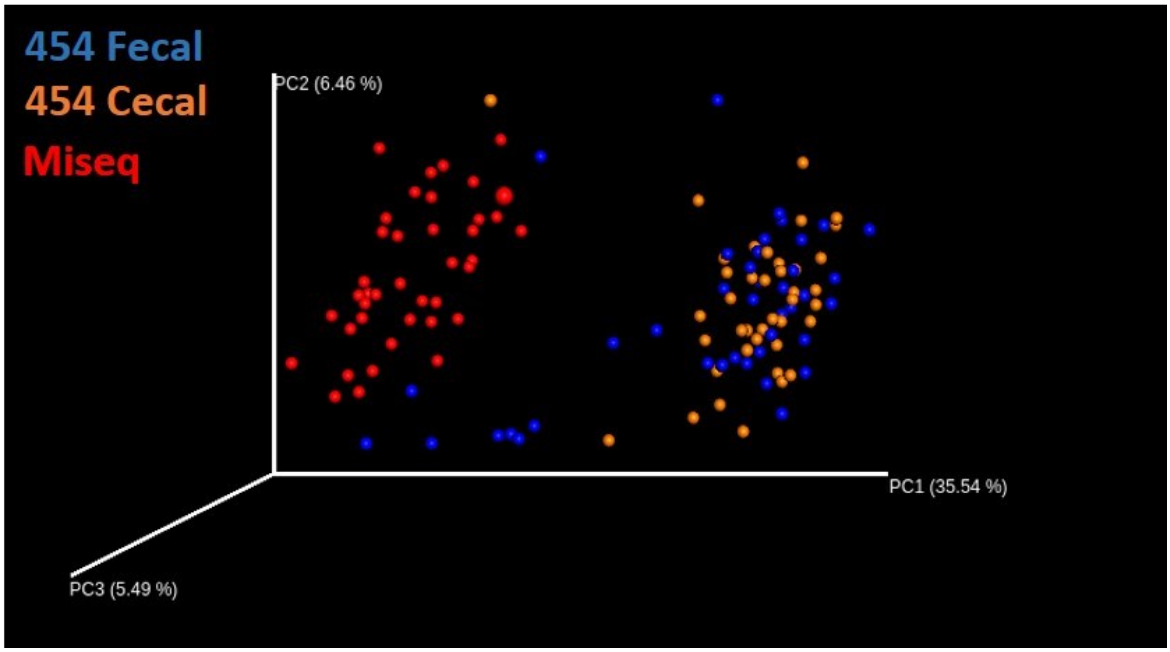

FIGURE 18B. PCoA of the weighted UniFrac values of all samples used in this experimentation.

In the PCoA analysis (Figure 18 A and B) there is no clear separation between the fecal and cecal data except for 11 fecal samples. The majority of the variation in the UniFrac estimates is explained PC1, which accounts for 35.54 % of the variation, followed by PC2 explains (6.46%), and PC3 explains (5.49%). This analysis shows that fecal and cecal samples have similar beta diversity.

Comparison of Core Microbiomes

Core microbiomes were identified for 95% of the samples in either of these datasets. In the fecal samples, the core biome consisted of 14 OTUs. Five of these were *Clostridia* and the remaining 9 were *Bacteroidetes*. Of the *Bacteroidetes*, 5 belonged to the family *Prevotellaceae* and 2 were in the family *Paraprevotellaceae*. The remaining 2 OTUs could not be identified to the family level. The 5 *Clostridia* OTUs included, 4 *Veillonellaceae* and only 1

Ruminococcaceae OTU (Table 7 and Table 17 in appendix). Species level identifications were not obtained.

TABLE 7. OTUs that comprise the core biome based on presence in 95% of the fecal samples.

454 Fecal Core Biome			
Phylum	Family	Genus	Count
Bacteroidetes	unknown	Unknown	2
	Paraprevotellaceae	Unknown	2
	Prevotellaceae	Prevotella	5
Firmicutes	Ruminococcaceae	Unknown	1
	Veillonellaceae	Phascolarctobacterium	4
total			14

The cecal samples had a much larger core biome than the fecal samples (Table 8 and Table 19 in Appendix). A total of 53 OTUs were identified as being a part of the core biome in the cecal data. Thirty-seven of these belonged to the class *Bacteroidia*; of these 10 in *Paraprevotellaceae*, 11 in family *Prevotellaceae*, 3 in *RF16*, and 13 unknown bacterial families. An additional 15 OTUs belonged to the class *Clostridia*. Of these; 3 were in *Lachnospiraceae*, 5 in *Ruminococcacea*, 5 *Veillonellaceae*, and 2 were not identified to the family level. There was 1 core OTU in the phyla *Verrucomicrobia* (Family *Verrucomicrobiales*).

TABLE 8. The OTUs that comprise the core biome and are present in 95% of the cecal samples.

454 Cecal Core Biom			
Phylum	Family	Genus	
Bacteroidetes	unknown	unknown	13
	Paraprevotellaceae	unknown	3
		Prevotella	3
		CF231	4
	Prevotellaceae	Prevotella	11
	RF16	unknown	3
Firmicutes	unknown	unknown	2
	Lachnospiraceae	unknown	2
		Roseburia	1
	Ruminococcaceae	unknown	3
		Ruminococcus	1
		flavefaciens	1
	Veillonellaceae	Anaerovibrio	2
		Phascolarctobacterium	3
Verrucomicrobia	Verrucomicrobiaceae	Akkermansia	1
Total			53

Identification of Shifts in GI Flora in Fecal and Cecal Samples

One of the major goals of laminitis research is to identify shifts in the hindgut microflora and then understand how these may contribute to the progression of the disease. Therefore, a critical aspect of these studies is understanding whether fecal and cecal samples show the same shifts in hindgut microflora in response to increase NSCs.

The 10 most commonly identified OTUs were tracked over the course of the 7-day treatment. The fecal samples exhibited a number of important bacterial OTUs increased in abundance based on the number of reads per sample (Figure 19 and Table 15 in appendix). Most notably, the *Paraprevotellaceae* OTUs increased from 10.36 % on the first day of treatment to 14.81% on the final day of the high carbohydrate treatment. ($P = 0.12$). *Prevotellaceae* also increased from 7.13% of to 14.12% of ($P = 0.09$). Finally, *Lachnospiraceae* OTUs increased

from 7.16% to 9.95% (P = 0.54). In contrast, *Ruminococcaceae* OTUs decreased in response to the high carbohydrate treatment from 8.35% to 4.01 (P = 0.02)

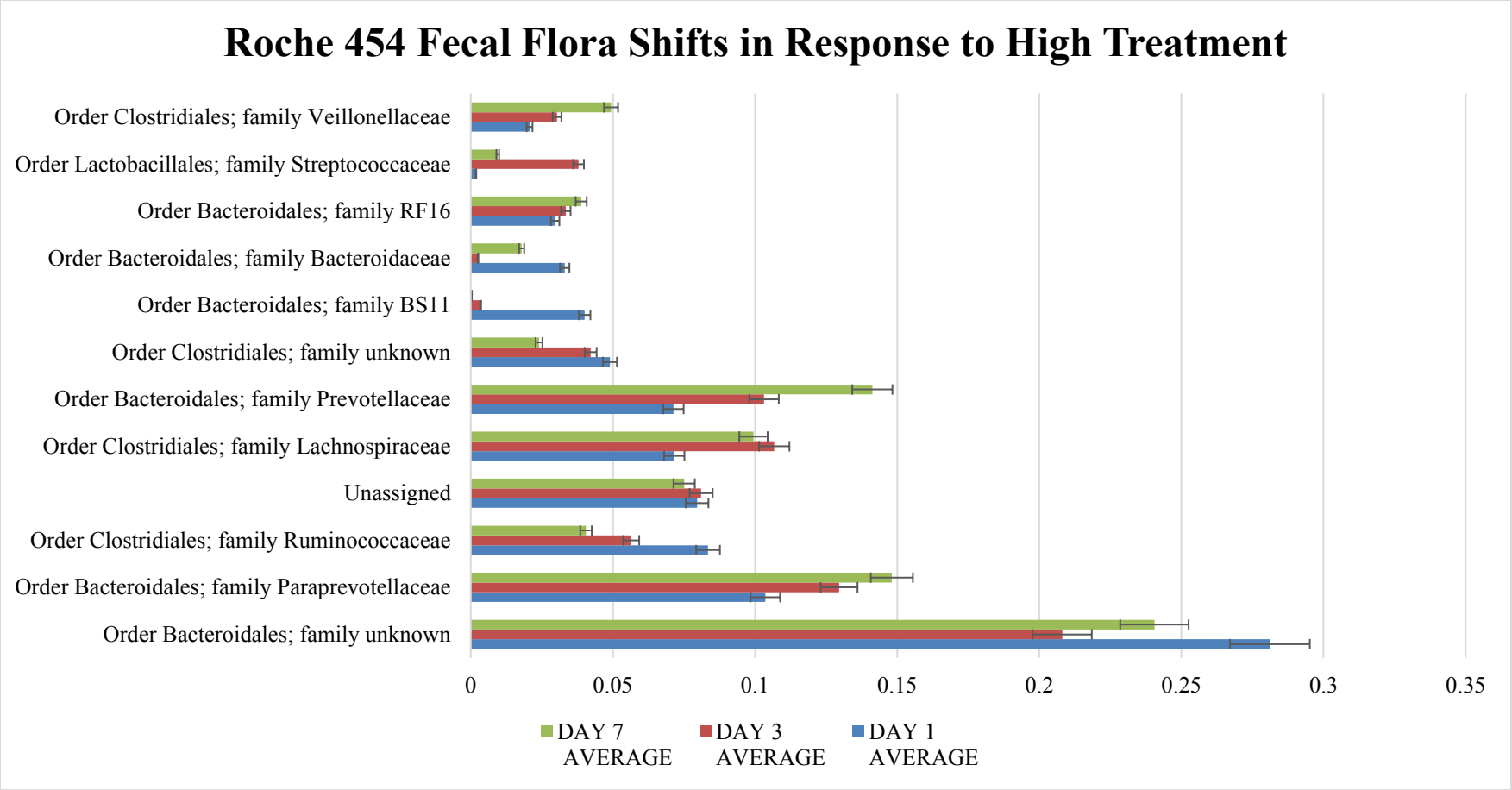

FIGURE 19. Percentage of total reads identified attributed to identified taxa in response to the high carbohydrate treatment over the course of 7 days. The error bars represent 95% CI.

There were notable differences in the cecal samples. Whereas the *Paraprevotellaceae* OTUs increased in frequency, a slight decrease was observed in the cecal samples (Figure 20 and Table 16 in appendix) from 18.61% to 17.03% (P=0.43). Similarly, the cecal dataset showed an increase in the number of *Prevotellaceae* reads, rising from 11.94% of the reads on the first day of the treatment to 16.01% of the reads on the last day of the high carbohydrate treatment (P = 0.13). *Ruminococcaceae* OTUs increased slightly over the 7 days of treatment, 4.41% of the total reads on day 1 and 5.46% of the total reads on day 7 (P= 0.49). In the fecal dataset, the *Ruminococcaceae* were shown to decrease in frequency over the 7 days of treatment. This change was not statistically significant (P=0.49)

Another noticeable increase was observed the family *Veillonellaceae*. On the first day of treatment, this family accounted for 1.48% of the total reads and rose to account for 6.04% of the total reads on day 7. This increase was also observed in the fecal dataset and is a statistically significant change in abundance (P=0.02). An unexplained spike in *Pasteurellaceae* was observed on day 3 but returned to near the percentage of reads that it occupied on day 1.

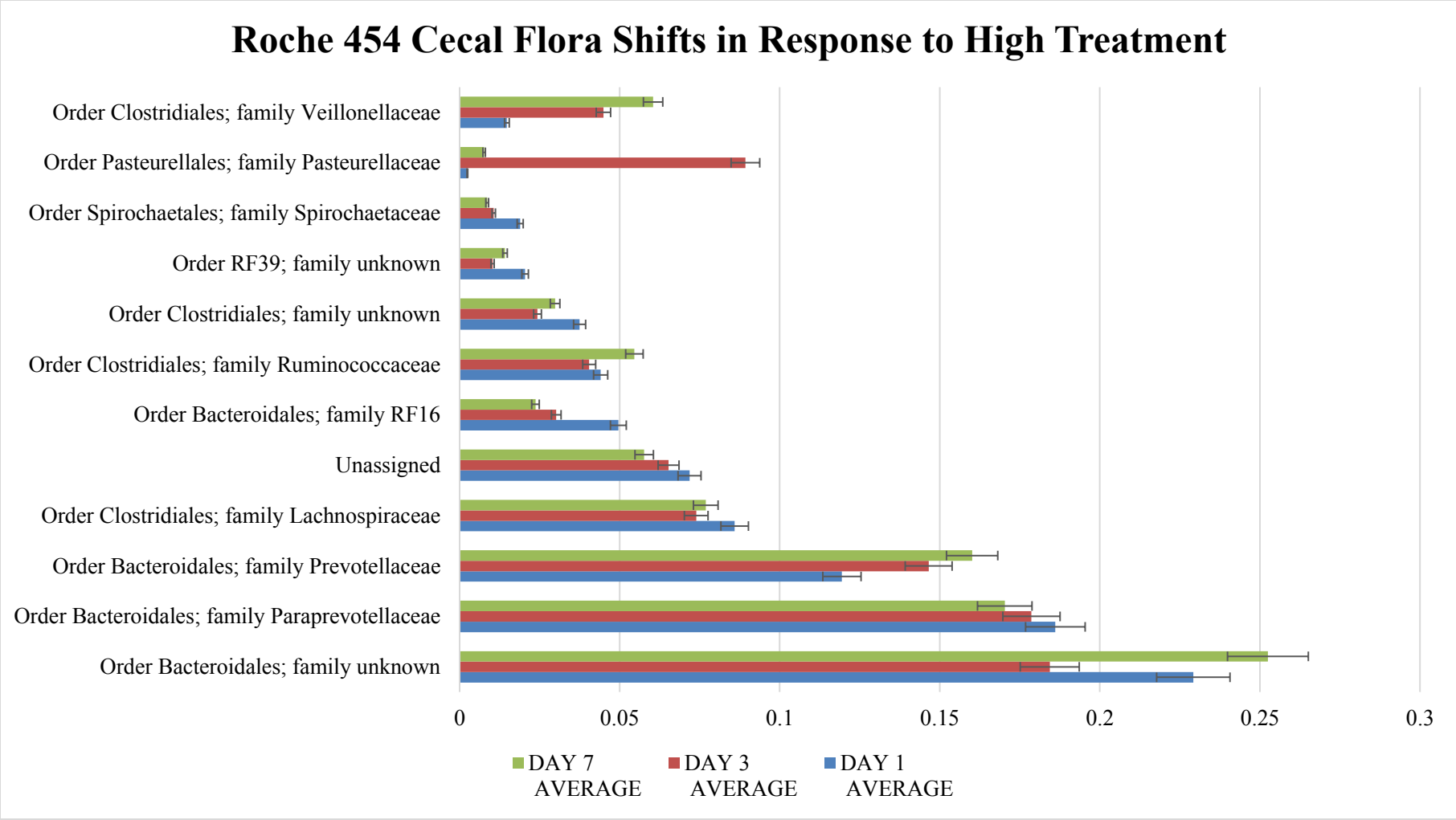

FIGURE 20. Visual representation of the shift in GI microflora in response to the high carbohydrate treatment. *Ruminococcaceae* and *Paraprevotellaceae* were shown to decrease. Additionally, *Prevotellaceae* and *Veillonellaceae* OTUs increased in frequency over the course of the treatment. The error bars represent 95% CI.

Chapter 5. Discussion and Conclusions

Comparison of Fecal Samples Sequenced with the Illumina MiSeq and the Roche 454 FLX

Comparing the Ability of Each Protocol to Identify OTUs

The first major conclusion is that the number of reads in the MiSeq method significantly increased the OTUs identified and, thus, the apparent makeup of the community. The MiSeq was able to produce over twice as many reads at only 1/4th the cost. Previous studies indicate that when a population, whether or not it is bacterial, is sampled, the number of species or OTUs will increase in positive correlation with the sampling effort (Hughes et al. 2001). More time and effort put into surveying results in a greater number of OTUs being identified, until all OTUs present are observed. It is challenging to observe all bacterial OTUs because some species are at low frequencies and others less readily amplified using universal 16S rRNA primers (Hughes et al. 2001; Acinas et al. 2005; Kanagawa 2003). Because the MiSeq generates more reads than the Roche 454 (~10 million per run compared to ~1 million, respectively) at less cost (approximately US$3,000 compared to US $12,000), studies can generate more reads per sample with the same budget. This means that it is much more likely obtain a complete view to the communities. This leads to improved estimates of OTU richness, *chao1,* and other diversity estimators

Just because the MiSeq produces more reads, does not mean that it is always necessary. There is PCR bias when amplifying bacteria (Acinas et al. 2005; Kanagawa 2003), therefore the additional reads may only provide more coverage of a specific subset of bacteria which are more easily amplified. In these cases, the additional reads may in reality be a hindrance, because they over represent certain OTUs. However, the data from this study, suggest that the additional

reads produced by the MiSeq significantly improve the ability to identify additional OTUs. The MiSeq was able to identify a significantly greater number of OTUs (i.e., higher OTU richness). Additionally, the *chao1* for the MiSeq fecal datasets was significantly greater than the 454 dataset. The *chao1* estimator takes into account the potential low abundance OTUs that may be missed by limited sampling.

$$Chao1 = S_{obs} + (n_1^2/n_2^2)$$

n1 represents OTUs that were only observed once and *n2* represents the OTUs that were observed multiple times (*n2*) (Chao 1984). Had the MiSeq sample only identified previously observed OTUs, the *chao1* value would not continue to rise with the addition of more reads because the identification of reads that had previously been recorded would cause the value for *n2* to increase. This would lead to a decrease in the correction factor added to the observed number of OTUs. At 10,000 randomly selected reads per sample the MiSeq is still identifying many novel OTUs and at a greater rate than what was observed when using the 454. This indicates that at 10,000 reads, the MiSeq was identifying novel OTUs more frequently than the 454. Interestingly enough, even though the curve for the 454 was shallower than that of the MiSeq, the 454 was still identifying new OTUs between 5,000 and 10,000 reads. Because the 454 protocol was still identifying novel OTUs at the extent of the reads produced in this study, it is likely that had more reads been sequenced with the 454 more novel OTUs would have been identified past 10,000 reads.

The Simpson's D and Shannon's H values also indicate that the extra reads sequenced by the MiSeq are novel. Both indicate that the relative evenness in is not significantly different from the 454 dataset. The rarefaction curves and the Simpson's D and Shannon's H values indicate that the additional MiSeq reads are identifying novel OTUs that the Roche 454 is

missing. Had the extra reads of the MiSeq identified previously, and more readily amplified, bacteria it would cause the MiSeq to be significantly more uneven than the 454 which was not observed (P=0.45, 0.39 and 0.47)

The Roche 454 protocol did not produce enough reads to identify all of the OTUs in horse fecal samples and that the data produced using the Illumina MiSeq is likely more comprehensive. However, this does not necessarily mean that the additional OTUs identified mediate to the response to higher ingestion of nonstructural carbohydrates or are biologically relevant to laminitis.

Comparing the Core Microbiomes Identified by Either Protocol

The core biome includes the dominant OTUs in a community and is therefore likely the most biologically relevant taxa to the horse. Similarly, to total OTUs, the MiSeq was able to identify a greater number of core bacteria than the 454, 104 compared to 14. This included one *Cyanobacteria* not present in the core OTUs for the 454 data; however, other *Cyanobacteria* at low abundance were observed in those samples. Initially, it was speculated that the presence of the *Cyanobacteria* was due to the contamination. However, C*yanobacteria* do comprise a small portion of the GI tract in mammals and are thought to be responsible for neurologic diseases such as amyotrophic lateral sclerosis, Alzheimer's disease, and Parkinson-Dementia-Complex in humans and Equine Motor Neuron Disease in Horses (Brenner 2013). Therefore, the presence of *Cyanobacteria* in the GI tract of the horses is not abnormal. Additionally, in the core biome the MiSeq was able to identify 2 OTUs in the order *TF7* and 4 in the order *Verrucomicrobia*, 4 OTUs were in the phyla *Spirochaetes*, and an additional 2 OTUs from the phylum *Proteobacteria*. These were not in the 454 core microbiome. The remaining 92 bacteria identified as core in the MiSeq dataset and all core OTUs in the 454 fecal dataset were either

Firmicutes or *Bacteroidetes*. The MiSeq reads had more observed *Firimcutes* species, with 41 distinct *Ruminococcaceae* OTUs indicating that this family is one of the most abundant and diverse lineage in the GI tract of horses. In contrast, only 1 *Ruminococcaceae* OTU was in the core microbe in the 454 dataset. Previous studies have also shown that *Ruminococcaceae* species are in the GI tract of horses are a part of the core microbiome, and are digestively significant (Steelman et al. 2012; Dougal et al. 2013; Costa et al. 2012). Therefore, the MiSeq protocol is better suited for detecting diverse OTUs and the core microbes in horses; however, the 454 protocol does appear to have one advantage; identify Gram-negative *Bacteroidetes* more readily than the MiSeq. Of the 104 core bacteria identified by the MiSeq only 21 were *Bacteroidetes*. This accounts for 20.2% of the core bacteria. The 454 identified 9 *Bacteroidetes* as core, which is 64.3% of the core biome. Not only did the 454 protocol identify a greater percentage of the core biome as being *Bacteroidetes* OTUs, but it was better suited to identify them. Of the 21 *Bacteroidetes* OTUs identified as being core OTUs in the MiSeq dataset, only 8 were identified to the family level. In contrast, 7 of the 9 *Bacteroidetes* core OTUs were assigned to a family in the 454 dataset. It appears as if the 454 FLX protocol was better able to identify a greater percentage of *Bacteroidetes* OTUs to family. However, the MiSeq protocol may be better suited for identifying *Firimcutes*. The identification of additional bacteria, as well as the greater numbers of both *Firimcutes* and *Bacteroidetes* OTUs may be the result of more reads produced by the MiSeq.

Comparing the Apparent Shifts in the Microflora in Response the High Carbohydrate Treatment

The changes in the GI flora over the 7 days of the high carbohydrate treatment also suggests that the MiSeq is better suited for the identification of Gram-positive *Firmicutes*. The results of this analysis were contradictory between the two different methods. In the MiSeq

protocol, there was an increase in *Ruminococcaceae*, however, in the 454 data this family decreased in abundance. Increases in Gram-negative *Bacteroidetes* were observed in both sample sets, but it was more substantial in the 454 dataset. Because the protocols were compared in the same DNA samples, the differences in the results is a function of the two protocols rather the samples or the bacterial communities. Therefore, the results suggest the MiSeq protocol is better suited for detecting shifts in *Firimcutes* and the 454 protocol is better suited for assaying *Bacteroidetes.*

The proposed pathology of laminitis states that as non-structural carbohydrates enter the hindgut, bacteria increase the rate of fermentation. This changes the pH and cause a shift in the GI microflora, such as increase in pH tolerant fermenting bacteria. The rise in *Ruminococcaceae* observed in the MiSeq data is consistent with this pathology. *Ruminococcaceae* are the bacteria primarily responsible for fermenting structural carbohydrates in the mammalian gut (Biddle et al. 2013). It is likely that the treatments were not severe enough to cause enough of a shift in bacteria to an extent which would to cause any significant damage to the mucosa. However, the high treatment did result in significantly lower pH of the cecal matter after each meal ($P \leq 0.01$) (Warzecha 2013). Additionally, the pathology proposes that due to the change in the pH of the gut, the mucosa is damaged and the gut wall becomes more permeable. Because of this increased permeability, endotoxins secreted by the cell wall in Gram-negative bacteria are absorbed, which leads to endotoxemia and laminitis. This suggests that both Gram-positive and Gram-negative bacteria are relevant in the pathology of carbohydrate induced laminitis. *Prevotella* species, specifically *Prevotella ruminicola*, has been observed to grow at low pH (Russel & Wilson 1996). *Prevotella* and the closely related family *Paraprevotella* were observed in all datasets, and an increase in these families would also be consistent with the

pathology of laminitis. Assuming the 454 protocol is more suited for identifying Gram-negative bacteria, the differences in the taxa observed between the 454 and the MiSeq is consistent with expectations. This suggests that the apparent biases in the 454 and MiSeq may be complementary as they can provide a more complete survey of different taxa of bacteria if they are combined in a study. This can explain the results observed in the taxa summaries and the shifts in GI flora. Over the course of the trial, the 454 protocol was more likely to preferentially amplify and survey Gram-negative bacteria; even though *Ruminococcaceae* OTUs were increasing the 454 protocol was preferentially amplifying *Prevotella and Paraprevotella*. In contrast, the MiSeq showed *Ruminococcaceae, Prevotella, and Paraprevotella* all increased but the Gram-negative bacteria did not exhibit as noticeable an increase. *Ruminococcaceae* is one of the most abundant bacterial families in the mammalian gut (Biddle et al. 2013) and this is consistent with the greater frequency of *Ruminococcaceae* compared to *Prevotella* and *Paraprevotella.* In contrast, the higher prevalence of *Paraprevotella* competed to *Ruminococcaceae,* in 454 data is not consistent with previous observations.

Conclusion

The outputs of the MiSeq dataset are more consistent with previous observations of the GI microflora of mammals as well as the proposed pathology of laminitis. The greater number of reads produced by the MiSeq increases its ability to identify OTUs compared to the 454. Due to its consistency with previous observations and the added power observed in this study, the Illumina MiSeq protocol was superior to that of the Roche 454 FLX for the identification of shifts in the gut bacteria using the 16s rRNA gene, particularly as it generates more data for less than ¼ of the cost. It is important to note that, had a researcher only been interested in the change in Gram-negative bacteria, then the 454 FLX protocol may be advisable. It is also likely

that the primary difference in the taxa summaries may be due to the use of different primers in the different protocols. Therefore, the use of the 454 primer set with the MiSeq may also provide the benefit of more reads along with the preferential amplification of *Bacteroidetes.* To further test this, cultured bacterial species for the major groups identified in this study should be isolated, and DNA should be extracted and quantified. The extracted DNA samples should be normalized and pooled so that there are equal amounts of DNA for all bacteria included. The pooled DNA samples should be amplified using both the primer sets used in the MiSeq and the 454 protocol and then sequenced on a MiSeq. Any difference observed could be attributed to the primer set.

Comparison of Fecal and Cecal Bacteria

Comparing the Alpha Diversity of Each Sample Dataset

The datasets were amplified and sequenced using the 454 protocol. Fecal and cecal samples were taken at the same time points from the same animal and DNA from both types of samples extracted with the same protocol. There was no significant difference between the number of OTUs that were identified in either sample dataset (P-value>0.05). Even though the OTU richness was not significantly different, the *chao1* value of the cecal samples was significantly higher than the fecal samples. This suggests that there are a greater number of taxa in the cecum than there was in the feces. The increase in the *chao1* is likely due to the identification of a greater number of singleton OTUs (*n1*). This causes the *chao1* to increase by increasing the corrective factor that is added to OTU richness. There are two explanations for this result. The first is that the number of OTUs present in the cecal samples is actually higher. The rarefaction curves using 10,000 randomly selected reads were consistent. No section of either graph exhibited an abnormal rise in the *chao1* value that was uncharacteristic with the rest

of the curve. This serves to show that the cecal sample dataset was not artificially higher due to the random amplification of more singleton OTUs than the fecal samples. In addition, cecal and fecal samples and similar evenness in their bacterial communities.

Comparison of the Beta Diversity between Fecal and Cecal Samples

The beta diversity observed between samples was higher in the fecal samples than it was in the cecal samples. More interestingly, based on the average weighted *UniFrac* values, the fecal and cecal samples exhibited less diversity between each them than there was among the fecal samples. This indicates greater variation in the communities observed in the fecal samples. This was supported by the PCoA, which showed the majority of fecal and cecal samples clustered together. However, a noticeable number of fecal samples clustered outside the main group. The cecal samples formed a tighter group in the PCoA analysis than the fecal samples. The PCoA also showed that the MiSeq fecal samples clustered separate from the 454 cecal and fecal samples. This indicated that the differences between the two sequencing methods contributes more to the observed diversity than the origin of the sample. The results indicate that there is less variation in the GI flora is of the cecum.

Comparison of Core Biomes

The most noticeable difference between the core biomes identified in the fecal and cecal samples was the number of bacteria. The fecal samples had 14 core OTUs whereas the cecal samples 53 core OTUs. Nine of the 14 core OTUs in the fecal samples were *Bacteroidetes* and 37 of the 53 cecal core OTUs were *Bacteroidetes*. This corresponds to 64.3% and 69.8% of the core biome, respectively. There were family level differences in this Phylum. Of the *Bacteroidetes,* 56% were *Prevotellaceae* in the fecal dataset, whereas 30% were *Prevotellaceae* in the cecal dataset. The percentage of *Bacteroidetes* that were *Paraprevotellaceae* was more

similar (22% and 27%, respectively). Additionally, the amount of *Firmicutes* OTUs identified as core were also comparable between the fecal and cecal samples. Approximately 35.7% of the fecal OTUs and 30.1% of the cecal OTUs were *Firmicutes.* The most common *Firmicutes* in the fecal dataset was *Veillonellaceae*. In the cecal dataset *Veillonellaceae* and *Ruminococcacea* were the most prevalent *Firmicutes.* In conclusion, the major difference between the core biomes of the 454 fecal and the 454 cecal samples was the number of core OTUs identified. Despite differences in the number of core OTUs, the core biomes had a similar composition and were substantially different from the MiSeq data.

Response to High Carbohydrate Treatment

There were some discrepancies in the microflora changes observed. Most notably, a decrease in *Paraprevotellaceae* OTUs and a slight increase in *Ruminococcaceae* OTUs was observed in the cecal dataset over the course of a 7-day high carbohydrate treatment. These were not observed in the fecal dataset. A substantial increase in *Paraprevotellaceae* was observed in addition to a substantial decrease in the percentage of reads identified as *Ruminococcaceae* in the feces. With the exception of these two taxa, the shifts identified in the GI flora were consistent across the two sample datasets. Taxa such as *Prevotellaceae* and *Veillonellaceae* were both observed to increase. Additionally, when discrepancies were identified, the results were only marginal. For example, whereas *Paraprevotellaceae* and *Lachnospiraceae* were identified as increasing in the fecal sample, there was not a noticeable change in the cecal samples.

The purpose of this study was to evaluate the common practice of using fecal samples to study the equine GI microflora and to describe the normal shift that occurs in healthy horses on a high carbohydrate diet. This was done to develop a foundation for studies that seek to understand bacterial factors in the GI tract of horses during the development of laminitis. The

goal of this study was not to induce laminitis in these horses. The use of a treatment that would have led to laminitis would have most likely provided clearer shifts in the GI microflora but would have also likely lead to the horses to being euthanized. Though the shifts may not be as substantial as they might have been if the treatments were more extreme, the results of this study can be extrapolated. It is likely the shifts would continue if the treatment was increased in magnitude or duration. The fecal samples show more substantial shifts in the GI microflora. This is evident in several families including P*araprevotellaceae, Ruminococcaceae,* and *Lachnospiraceae*, whereas in the cecal dataset these families change minimally.

This observation may be explained the constant movement of the chyme in the horse's intestines and thus the products of fermentation in the cecum are constantly moving out of the compartment. Therefore, even though the lactic acid and the VFAs proposed to induce laminitis are first produced in the cecum, they continue to travel with the chyme into the downstream portions of the hindgut where they can accumulate and have a greater effect on downstream portions of the large intestine. This hypothesis is supported by the observation that when laminitis is induced via carbohydrate overload in horses, one of the earliest signs of intestinal distress is the presence of acidic diarrhea (Van den Berg et al. 2012; Pollitt 2004). Observing greater changes in the microflora of the feces rather than the cecal material is consistent with the digestive process in the horse. In closing, when the goal is to observe shifts in the GI flora in laminitic horses, the data presented in this study suggest clearer trends may be observed in feces. Therefore, feces appear to be are a better candidate for studying hindgut microflora than the cecal matter.

Conclusions

The data presented in this study has led to three main conclusions.

1. The increased number of reads generated by the Illumina MiSeq adds additional power and to observed OTUs in the GI microflora when compared to Roche 454 FLX.
2. The sequencing protocols used in this study had a greater effect on the observed microbial diversity than the origin of the sample. This is likely due to both differences in primers and amount of reads per sample. The results from the Illumina protocol appears is more informative for bacteria that may be implicated in the pathology of carbohydrate induced laminitis. Additionally, the primers used in the Roche protocol seem to exhibit a bias towards Gram-negative bacteria when compared to the results of the Illumina protocol.
3. Fecal microflora exhibited a more substantial shift than the cecal microflora in response to high carbohydrate treatment. It is hypothesized that this is due to the migration of lactic acid and VFAs out of the cecum with the chyme to downstream portion of the GI tract were it accumulates.

Works Cited

Acinas, S. G., Sarma-Rupavtarm, R., Klepac-Ceraj, V., & Polz, M. F. (2005). PCR-induced sequence artifacts and bias: Insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Applied and Environmental Microbiology*, *71*(12), 8966-8969.

Al Jassim, R., & Andrews, F. (2009). The bacterial community of the horse gastrointestinal tract and its relation to fermentative acidosis, laminitis, colic, and stomach ulcers. *Veterinary Clinics; Equine Practices*, *25*(2), 199-215.

Al Jassim, R., Scott, P. T., Trebbin, A. L., Trott, D., & Pollitt, C. C. (2005). The genetic diversity of lactic acid producing bacteria in the equine gastrointestinal tract. *FEMS Microbiology Letters*, *248*, 75-81.

Al Jassim, R. A., & Rowe, J. B. (1999). Better understanding of acidosis and its control. *Recent Advances in Animal Nutrition in Australia*, *12*, 91-97.

Asplin, K. E., Sillence, M. N., Pollitt, C. C., & McGowan, C. M. (2007). Induction of laminitis by prolonged hyperinsulinaemia in clinically normal ponies. *The Veterinary Journal*, *174*(3), 530-535.

Baileya, S. R., Adairb, H. S., Reinemeyerc, C. R., Morgand, S. J., Brookse, A. C., Longhoferf, S. L., & Elliott, J. J. (2009). Plasma concentrations of endotoxin and platelet activation in the developmental stage of oligofructose-induced laminitis. *Veterinary Immunology and Immunopathology*, *129*(3-4), 167-173.

Baselga, A., & Orme, C. D. (2012). Betapart: An R package for the study of beta diversity. *Methods in Ecology and Evolution*, *3*, 808-812.

Biddle, A., Stewart, L., Blanchard, J., & Leschine, S. (2013). Untangling the genetic basis of fibrolytic specialization by lachnospiraceae and ruminococcaceae in diverse gut communities. *Diversity*, *5*(3), 627-640.

Block, J. A., & Sequeira, W. (2001). Raynaud's phenomenon. *The Lancet*, *357*(9273), 2042-2048.

Brenner, S. R. (2013). Blue-green algae or cyanobacteria in the intestinal micro-flora may produce neurotoxins such as Beta-N-Methylamino-l-Alanine (BMAA) which may be related to development of amyotrophic lateral sclerosis, Alzheimer's disease and Parkinson-Dementia-Complex in humans and Equine Motor Neuron Disease in Horses. *Medical Hypotheses*, *80*(1), 103.

Caporaso, J. G., Bittinger, K., Bushman, F. D., DeSantis, T. Z., Andersen, G. L., & Knight, R. (2010). PyNAST: A flexible tool for aligning sequences to a template alignment. *Bioinformatics*, *26*(2), 266-267.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., . . . Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, *7*, 335-336.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., . . . Knight, R. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The International Society for Microbial Ecology Journal*, *6*(8), 1621-1624.

Chao, A. (1984). Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics*, *11*(4), 265-270.

Chen, J., Bittinger, K., Charlson, E. S., Hoffmann, C., Lewis, J., Wu, G. D., . . . Li, H. (2012). Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics*, *28*(16), 2106-2113.

Costa, M. C., Arroyo, L. G., Allen-Vercoe, E., Stämpfli, H. R., Kim, P. T., Sturgeon, A., & Weese, J. S. (2012). Comparison of the fecal microbiota of healthy horses and horses with colitis by high throughput sequencing of the V3-V5 region of the 16S rRNA gene. *PLOS One*. doi:10.1371/journal.pone.0041484

Cummings, J. H. (1981). Short chain fatty acids in the human colon. *Gut*, *22*(9), 763-779.

Daly, K., Stewart, C. S., Flint, H. J., & Shirazi-Beechey, S. P. (2001). Bacterial diversity within the equine large intestine as revealed by molecular analysis of cloned 16S rRNA genes. *FEMS Microbial Ecology*, *38*(2-3), 141-151.

DeJong, T. M. (1975). A comparison of three diversity indices based on their components of richness and evenness. *Oikos*, *26*(2), 222-227.

Dougal, K., De La Fuente, G., Harris, P. A., Girdwood, S. E., Pinloche, E., & Newbold, C. J. (2013). Identification of a core bacterial community within the large intestine of the horse. *PLOS One*. doi:10.1371/journal.pone.0077660

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460-2461.

Edwards, C. A., & Parrett, A. M. (2012). Intestinal flora during the first months of life: New perspectives. *British Journal of Nutrition*, *88*(1), S11-S18.

French, K. R., & Pollitt, C. C. (2004). Equine laminitis: Glucose deprivation and MMP activation induce dermo-epidermal separation in vitro. *Equine Veterinary Journal*, *36*(3), 261-266.

Garner, H. E., Coffman, J. R., Hahn, A. W., Hutcheson, D. P., & Tumbleson, M. E. (1975). Equine laminitis of alimentary origin: An experimental model. *American Journal of Vetrinary Research*, *36*(4), 441-444.

Gibbs, P. G., Potter, G. D., Schelling, G. T., Kreider, J. L., & Boyd, C. L. (1987). Digestion of hay protein in different segments of the equine digestive tract. *Journal of Animal Science*, *6*(2), 400-406.

Gibbs, P. G., Potter, G. D., Schelling, G. T., Kreider, J. L., & Boyd, C. L. (1996). The significance of small vs large intestinal digestion of cereal grain and oilseed protein in the equine. *Journal of Equine Veterinary Science*, *16*(2), 60-65.

Gomis-Ruth, F.-X., Maskos, K., Huber, R., Maskos, K., Betz, M., Bergner, A., . . . Bode, W. (1997). Mechanism of inhibition of the human matrix metalloproteinase stromelysin-1 by TIMP-1. *Letters to Nature*, *389*, 77-80.

Grubb, J. A., & Dehority, B. A. (1975). Effects of an abrupt change in ration from all roughage to high concentrate upon rumen microbial numbers in sheep. *Applied Microbiology*, *30*(3), 404-412.

Hamady, M., & Knight, R. (2009). Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Research*, *19*(7), 1141-1152.

Hood, D. M., Amoss, M. S., & Grosenbaugh, D. A. (1990). Equine laminitis; A potential model of Raynaud's phenomenon. *Angiology; The Journal of Vascular Diseases*, *41*(4), 270-277.

Hood, D. M., Grosenbaugh, D. A., Mostafa, M. B., Morgan, S. J., & Thomas, B. C. (1993). The role of vascular mechanisms in the development of acute equine laminitis. *Journal of Veterinary Internal Medicine*, *7*(4), 228-234.

How to calculate biodiversity [Pamphlet]. (n.d.). Gainsville, FL: University of Florida.

Hughes, J. B., Hellmann, J. J., Ricketts, T. H., & Bohannan, B. J. (2001). Counting the uncountable: Statistical approaches to estimating microbial diversity. *Applied and Environmental Microbiology*, *67*(10), 4399-4406.

Johnson, P. J., Tyagi, S. C., Katwa, L. C., Ganjam, V. K., Moore, L. A., Kreeger, J. M., & Messer, N. T. (1998). Activation of extracellular matrix metalloproteinases in equine laminitis. *The Veterinary Record*, *142*, 392-396.

Kanagawa, T. (2003). Bias and artifacts in multitemplate polymerase chain reactions (PCR). *Journal of Bioscience and Bioengineering*, *96*(4), 317-323.

Kern, D. L., Slyter, L. L., Leffel, E. C., Weaver, J. M., & Oltjen, R. R. (1974). Ponies vs. steers: Microbial and chemical characteristics of intestinal ingesta. *The Journal of Animal Science*, *38*(3), 559-564.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., & Glöckner, F. O. (2012). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acid Research, 41*(1).

Koleff, P., Gaston, K. J., & Lennon, J. J. (2003). Measuring beta diversity for presence–absence data. *Journal of Animal Ecology*, *72*, 367-382.

Kronfeld, D. S., Treiber, K. H., Hess, T. M., Splan, R. K., Byrd, B. M., Staniar, W. B., & Whitey, N. W. (2006). Metabolic syndrome in healthy ponies facilitates nutritional countermeasures against pasture laminitis. *The Journal of Nutrition*, *136*(7), 290-293.

Krueger, A. S., Kinder, D. A., Garner, H. E., & Sprouse, R. F. (1986). Ultrastructural study of the equine cecum during onset of laminitis. *American Journal of Veterinary Research*, *47*(8), 1804-1812.

Kyaw-Tanner, M., & Pollitt, C. C. (2004). Equine laminitis: Increased transcription of matrix metalloproteinase-2 (MMP-2) occurs during the developmental phase. *Equine Veterinary Journal*, *36*(3), 221-225.

Kyaw-Tanner, M. T., Wattle, O., Van Epps, A. W., & Pollitt, C. C. (2008). Equine laminitis: Membrane type matrix metalloproteinase-1 (MMP-14) is involved in acute phase onset. *Equine Veterinary Journal*, *40*(5), 482-487.

Latham, M. J., & Wolin, M. J. (1977). Fermentation of cellulose by ruminococcus flavefaciens in the presence and absence of methanobacterium ruminantium. *Applied and Environmental Microbiology*, *34*(3), 297-301.

Levin, J., Poore, T. E., Zauber, N. P., & Oser, R. S. (1970). Detection of endotoxin in the blood of patients with sepsis due to gram-negative bacteria. *New England Journal of Medicine*, *283*, 1313-1316.

Ley, R. E., Hamady, M., Lozupone, C., Turnbaugh, P., Ramey, R. R., Bircher, J. S., . . . Gordon, J. I. (n.d.). Evolution of mammals and their gut microbes. *Science*, *320*(5883), 1647-1651.

Lindberg, R. L., De Groot, C. J., Montagne, L., Freitag, P., Van der Valk, P., Kappos, L., & Leppert, D. (2001). The expression profile of matrix metalloproteinases (MMPs) and their inhibitors (TIMPs) in lesions and normal appearing white matter of multiple sclerosis. *Brain*, *124*(Pt. 9), 1743-1753.

Lloyd, L. E., Peckham, H. E., & Crampton, E. W. (1956). The effect of change of ration on the required length of preliminary feeding period in digestion trials with sheep. *The Journal of Animal Science*, *15*(3), 846-853.

Longland, A. C., & Byrdy, B. M. (2006). Pasteural non-structural carbohydrates and equine laminitis. *The Journal of Nutrition*, *136*(Suppl. 7), 2099-2102.

Lozupone, C., & Knight, R. (2005). UniFrac: A new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*, *71*(12), 8228-8235.

Mackie, R. I., & Wilkins, C. A. (1998). Enumeration of anaerobic bacterial microflora of the equine gastrointestinal tract. *Applied and Environmental Microbiology*, *54*(9), 2155-2160.

Macpherson, A. J., & Harris, N. L. (2004). Interactions between commensal intestinal bacteria and the immune system. *Nature Review Immunology*, *4*, 478-485.

Mallett, A. K., Bearne, C. A., & Rowland, I. R. (1983). Metabolic activity and enzyme induction in rat fecal microflora maintained in continuous culture. *The Journal of Applied and Environmental Microbiology*, *46*(3), 591-595.

Manefield, M., Whiteley, A. S., Griffiths, R. I., & Bailey, M. J. (2002). RNA stable isotope probing, a novel means of linking microbial community function to phylogeny. *Applied and Environmental Biology*, *68*(11), 5367-5373.

Mariat, D., Firmesse, O., Levenez, F., Guimar?es, V. D., Soko, H., Doré, J., . . . Furet, J. P. (2009). The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age. *BMC Microbiology*, *9*(123). doi:10.1186/1471-2180-9-123

McCune, B., Grace, J. B., & Urban, D. L. (n.d.). *Analysis of ecological communities* (2nd ed.). Gleneden Beach, OR: MJM Software Design.

Milinovich, G. J., Trott, D. J., Burrell, P. C., Van Eps, A. W., Thoefner, M. B., Blackall, L. L., . . . Pollitt, C. C. (2006). Changes in equine hindgut bacterial populations during oligofructose-induced laminitis. *Environmental Microbiology*, *8*(5), 885-898.

Mungall, B. A., Kyaw-Tanner, M., & Pollitt, C. C. (2001). In vitro evidence for a bacterial pathogenesis of equine laminitis. *The Journal of Veterinary Microbiology*, *79*(3), 209-223.

Navas-Molina, J. A., Peralta-Sánchez, J. M., González, A., McMurdie, P. J., Vázquez-Baeza, Y., Xu, Z., . . . Knight, R. (2013). Advancing our understanding of the human microbiome using QIIME. *Methods in Enzymology*, *531*, 371-444.

Newby, A. C. (2007). Metalloproteinases and vulnerable atherosclerotic plaques. *Trends in Cardiovascular Medicine*, *17*(8), 253-258.

Nocek, J. E. (1997). Bovine acidosis: Implications on laminitis. *Journal of Dairy Science*, *80*, 1005-1028.

Pollitt, C. C. (2004). Equine lamanitis. *Clinical Techniques in Equine Practice*, *3*(1), 34-44.

Pollitt, C. C. (2008). *Equine laminitis; Current concepts*. Canberra, Australia: Rural Industry Research and Development.

Pollitt, C. C., & Daradka, M. (1998). Equine laminitis basement membrane pathology: Loss of type IV collagen, type VII collagen and laminin immunostaining. *The Equine Veterinary Journal*, *26*, 139-144.

Pollitt, C. C., & Davies, C. T. (1998). Equine laminitis; Its development coincides with increased sublamellar blood flow. *Equine Veterinary Journal*, *30*(26), 125-132.

Pollitt, C. C., Pass, M. A., & Pollitt, S. (1998). Batimastat (BB-94) inhibits matrix metalloproteinases of equine laminitis. *Equine Veterinary Journal*, *30*(26), 119-124.

Pollitt, C. C., & Van Epps, A. W. (2006). Equine lamanitis incudced with fructose. *Equine Veterinary Journal*, *38*(3), 203-208.

Price, M. N., Dehal, P. S., & Arkin, A. P. (2009). FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular Biology and Evolution*, *26*(7), 1641-1650.

Ralston, S. L. (2005). Feeding dentally challenged horses. *Clinical Techniques in Equine Practice*, *4*(2), 117-119.

Romano, A. H., & Conway, T. (1996). Evolution of carbohydrate metabolic pathways. *Research in Microbiology*, *147*(6), 448-455.

Rotger, A., Ferret, A., Manteca, X., Ruiz de la Torre, J. L., & Calsamiglia, S. (2005). Effects of dietary nonstructural carbohydrates and protein sources on feeding behavior of tethered heifers fed high-concentrate diets. *Journal of Animal Science*, *84*(5), 1197-1204.

Schloss, P. D., & Westcott, S. L. (2011). Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. *Applied and Environmental Microbiology*, *77*(10), 3219-3226.

Sepkoski, J. J., Jr. (1988). Alpha, beta, or gamma: Where does all the diversity go? *Paeleobiology*, *14*(3), 221-234.

Shade, A., & Handelsman, J. (2012). Beyond the Venn diagram: The hunt for a core microbiome. *Environmental Microbiology*, *14*(1), 1-8.

Simberloff, D. (1972). Properties of the rarefaction diversity measurement. *The American Naturalist*, *106*(949), 414-418.

Sjaastad, O. V., Sand, O., & Hove, K. (2003). *Physiology of domestic animals*. Oslo, Norway: Scandinavian Veterinary Press.

Steelman, S. M., Chowdhary, B. P., Dowd, S., Suchodolski, J., & Janecka, J. E. (2012). Pyrosequencing of 16S rRNA genes in fecal samples reveals high diversity of hindgut

microflora in horses and potential links to chronic laminitis. *BMC Veterinary Research*, *231*(8). doi:10.1186/1746-6148-8-231

Tay, S. T., Ivanov, V., Yi, S., Zhuang, W. Q., & Tay, J. H. (2002). Presence of anaerobic bacteroides in aerobically grown microbial granules. *Microbial Ecology*, *44*(3), 278-285.

Toupadakis, C. A., Wong, A., Gentos, D. C., Cheung, W. K., Borjesson, D. L., Ferraro, G. L., . . . Yellowley, C. E. (2010). Comparison of the osteogenic potential of equine mesenchymal stem cells from bone marrow, adipose tissue, umbilical cord blood, and umbilical cord tissue. *American Journal of Veterinary Research*, *71*(10), 1237-1245.

Turnbaugh, P. J., Ley, R. E., Hamady, M. M., Fraser-Liggett, C. M., Knight, R. R., & Gordon, J. I. (2007). The human microbiome project. *Nature*, *449*(7164), 804-810.

Van Den Berg, M., Hoskin, S. O., Rogers, C. W., & Grinsberg, A. (2013). Fecal pH and microbial populations in thoroughbred horses during transition from pasture to concentrate feeding. *Journal of Equine Veterinary Science*, *33*(4), 215-222.

Van Epps, A., Collins, S. M., & Pollitt, C. C. (2010). Supporting limb laminitis. *Veterinary Clinics of North America: Equine Practice*, *28*(2), 287-302.

Vernon, R. G., & Flint, D. J. (1988). Lipid metabolism in farm animals. *Proceedings of the Nutrition Society*, *47*(3), 287-293.

Videnska, P., Faldynova, M., Juricova, H., Babak, V., Sisak, F., Havlickova, H., & Rychlik, I. (2013). Chicken faecal microbiota and disturbances induced by single or repeated therapy with tetracycline and streptomycin. *BMC Veterinary Research*, *9*(30). doi:10.1186/1746-6148-9-30

White, N. A. (2005). Equine laminitis. Paper presented at the Waltham Seminar, Washington DC.

Youssef, N. H., Couger, M. B., McCully, A. L., Criado, A. E., & Elshahed, M. S. (2015). Assessing the global phylum level diversity within the bacterial domain: A review. *Journal of Advanced Research*, *6*(3), 269–282.

Appendix

TABLE 9. Samples used in this analysis grouped by sequencing platform and the number of reads.

Sample Fecal MiSeq	Number of Reads	Sample 454 Fecal	Number of Reads
mj1	5399	1.f	10114
mj2	16637	2.f	6558
mj3	12777	3.f	9047
mj4	9479	4.f	9203
mj5	15301	5.f	14678
mj6	1889	6.f	3237
mj7	7376	7.f	10590
mj43	16989	43.f	6177
mj44	28484	44.f	6995
mj45	30472	45.f	11270
mj46	20941	46.f	10530
mj47	7862	47.f	13401
mj48	29665	48.f	5694
mj49	27771	49.f	6341
mj50	16917	50.f	9580
mj51	33940	51.f	11470
mj52	35767	52.f	10770
mj53	19811	53.f	11622
mj54	11586	54.f	12904
mj55	12086	55.f	9213
mj56	13734	56.f	7574
mj57	6000	58.f	6764
mj58	9409	59.f	11659
mj59	9944	60.f	8234
mj62	8157	61.f	6375
mj63	8364	62.f	6209
mj99	9487	63.f	9051
mj100	739	100.f	12335
mj101	13365	101.f	9651
mj102	14956	102.f	6983
mj103	21217	103.f	7695
mj104	39084	104.f	14624
mj105	23851	105.f	4454
mj106	24861	106.f	3866

mj107	27142	107.f	4775
mj108	27788	108.f	6577
mj109	26067	109.f	10610
mj110	21439	110.f	6591
mj111	142894	111.f	9317
mj112	17495	112.f	8423

TABLE 10. Weighted Unifrac Values

Weighted Unifrac Values			
Key	Description	Mean	St Dev
cec x cec	Mean diversity of Roche cecal samples	0.284762	0.111551
fec x fec	Mean diversity of Roche fecal samples	0.36298	0.14935
mis x mis	Mean diversity of MiSeq fecal samples	0.324964	0.163768
mis x cec	Mean diversity between Roche cecal and MiSeq fecal samples	0.58721	0.08701
mis x fec	Mean diversity between Roche fecal and MiSeq fecal samples	0.55249	0.104186
cec x fec	Mean diversity between Roche cecal and fecal samples	0.33327	0.14224
allfec x cec	Mean diversity between cecal and fecal samples	0.460239	0.173275
mis x roc	Mean diversity between Roche and MiSeq samples	0.56266	0.10832

TABLE 11. Unweighted Unifrac Values

Unweighted Unifrac Values			
Key	Description	Mean	St Dev
cec x cec	Mean diversity of all Roche cecal samples	0.657298	0.115671
fec x fec	Mean diversity of all Roche fecal samples	0.71619	0.13217
mis x mis	Mean diversity of all MiSeq fecal samples	0.698923	0.122848
mis x cec	Mean diversity between Roche cecal and MiSeq fecal samples	0.88552	0.01869
mis x fec	Mean diversity between Roche fecal and MiSeq fecal samples	0.877161	0.028836
cec x fec	Mean diversity between Roche cecal and fecal samples	0.71182	0.06238
allfecx cec	Mean diversity between Roche cecal and all fecal samples	0.818351	0.080021
mis x roc	Mean diversity between Roche and MiSeq samples	0.88134	0.02465

TABLE 12. Frequency of 10 most abundant OTUs during a 7-day high carbohydrate treatment in the fecal samples sequenced with the MiSeq.

Shifts in the Fecal Flora Identified by the Miseq in Response to the High Treatment			
#OTU ID	DAY 1 AVERAGE	DAY 3 AVERAGE	DAY 7 AVERAGE
Unassigned	25.09%	12.78%	5.28%
Order Bacteroidales; family unknown	15.89%	18.27%	16.65%
Order Clostridiales; family Ruminococcaceae	13.15%	15.69%	20.27%
Order Clostridiales; family unknown	6.78%	5.95%	8.25%
Order Clostridiales; family Lachnospiraceae	5.72%	8.12%	7.63%
Order Spirochaetales; family Sipirochaetaceae	4.31%	4.83%	4.34%
Order Bacteroidales; family Paraprevotellaceae	3.77%	5.15%	5.13%
Order Fibrobacterales; family Fibrobacteraceae	2.80%	1.78%	2.10%
Order CW040; family F16	2.10%	1.39%	2.98%
Order WCHB1-41; family RFP12	1.80%	1.43%	1.87%
Order Clostridiales; family Veillonellaceae	1.41%	2.87%	2.63%
Order Bacteroidales; family BS11	1.11%	2.63%	2.15%
Order Bacteroidales; family Prevotellaceae	0.94%	2.26%	1.81%
Order Clostridiales; family Clostridiaceae	0.85%	1.27%	2.47%

TABLE 13. Frequency of 10 most abundant OTUs during a 7-day high carbohydrate treatment using 454.

Shifts in the Fecal Flora Identified in Using the 454 in Response to the High Treatment			
#OTU ID	DAY 1 AVERAGE	DAY 3 AVERAGE	DAY 7 AVERAGE
Order Bacteroidales; family unknown	28.11%	20.81%	24.05%
Order Bacteroidales; family Paraprevotellaceae	10.36%	12.95%	14.81%
Order Clostridiales; family Ruminococcaceae	8.35%	5.64%	4.05%
Unassigned	7.96%	8.11%	7.51%
Order Clostridiales; family Lachnospiraceae	7.16%	10.68%	9.95%
Order Bacteroidales; family Prevotellaceae	7.13%	10.32%	14.13%
Order Clostridiales; family unknown	4.90%	4.22%	2.40%
Order Bacteroidales; family BS11	4.01%	0.35%	0.04%
Order Bacteroidales; family Bacteroidaceae	3.31%	0.26%	1.79%
Order Bacteroidales; family RF16	2.97%	3.34%	3.89%
Order Lactobacillales family Streptococcaceae	0.19%	3.79%	0.95%
Order Clostridiales; family Veillonellaceae	2.07%	3.04%	4.94%

TABLE 14. Total number of reads generated using the 454 method for each sample. Samples were taken at the same time points and for the same horses,

Fecal 454 Samples	# of Reads Identified	Cecal 454 Samples	# of Reads Identified
1.f	10114	1	7202
2.f	6558	2	7986
3.f	9047	3	9449
4.f	9203	4	7985
5.f	14678	5	8267
6.f	3237	6	7798
7.f	10590	7	7005
43.f	6177	43	10933
44.f	6995	44	8207
45.f	11270	45	7588
46.f	10530	46	8147
47.f	13401	47	6607
48.f	5694	48	7799
49.f	6341	49	2154
50.f	9580	50	7362
51.f	11470	51	5079
52.f	10770	52	7396
53.f	11622	53	9290
54.f	12904	54	7503
55.f	9213	55	9715
56.f	7574	56	8193
58.f	6764	58	7841
59.f	11659	59	7412
60.f	8234	60	8014
61.f	6375	61	5639
62.f	6209	62	6859
63.f	9051	63	7247
100.f	12335	100	7983
101.f	9651	101	9186
102.f	6983	102	6047
103.f	7695	103	12629
104.f	14624	104	10607
105.f	4454	105	9081
106.f	3866	106	7219
107.f	4775	107	36349

108.f	6577	108	8773
109.f	10610	109	7502
110.f	6591	110	6284
111.f	9317	111	7216
112.f	8423	112	6319

TABLE 15. Changes in the hindgut microflora in response the high carbohydrate treatment based on frequency of total number of reads.

Roche 454 Fecal Flora Shifts in Response to High Treatment			
#OTU ID	DAY 1 AVERAGE	DAY 3 AVERAGE	DAY 7 AVERAGE
Order Bacteroidales; family unknown__	28.11%	20.81%	24.05%
Order Bacteroidales; family Paraprevotellaceae	10.36%	12.96%	14.81%
Order Clostridiales; family Ruminococcaceae	8.35%	5.64%	4.05%
Unassigned	7.96%	8.11%	7.51%
Order Clostridiales; family Lachnospiraceae	7.14%	10.68%	9.95%
Order Bacteroidales; family Prevotellaceae	7.13%	10.32%	14.13%
Order Clostridiales; family unknown	4.9%	4.22%	2.40%
Order Bacteroidales; family BS11	4.01%	0.35%	0.04%
Order Bacteroidales; family Bacteroidaceae	3.31%	0.26%	1.79%
Order Bacteroidales; family RF16	2.97%	3.34%	3.89%
Order Lactobacillales; family Streptococcaceae	0.19%	3.79%	0.95%
Order Clostridiales; family Veillonellaceae	2.07%	3.04%	4.94%

TABLE 16. Reports the changes in the cecal microflora in response the high carbohydrate treatment. Samples were sequenced with the Roche 454 FLX high throughput sequencer.

Roche 454 Cecal Flora Shifts in Response to High Treatment			
#OTU ID	DAY 1 AVERAGE	DAY 3 AVERAGE	DAY 7 AVERAGE
Order Bacteroidales; family unknown	22.92%	18.43%	25.25%
Order Bacteroidales; family Paraprevotellaceae	18.61%	17.86%	17.023%
Order Bacteroidales; family Prevotellaceae	11.94%	14.65%	16.01%
Order Clostridiales; family Lachnospiraceae	8.59%	7.39%	7.69%
Unassigned	7.18%	6.53%	5.77%
Order Bacteroidales; family RF16	4.96%	3.02%	2.37%
Order Clostridiales; family Ruminococcaceae	4.41%	4.05%	5.46%
Order Clostridiales; family unknown	3.75%	2.43%	2.98%
Order RF39; family unknown	2.05%	1.03%	1.42%
Order Spirochaetales; family Spirochaetaceae	1.89%	1.07%	0.86%
Order Pasteurellales; family Pasteurellaceae	0.24%	8.93%	0.77%
Order Clostridiales; family Veillonellaceae	1.48%	4.49%	6.04%

TABLE 17. Core OTUs identified in 95% of the fecal samples sequenced with the Roche 454.

454 Fecal Core Biome				
OTU	Phylum	Class	Order	Family
denovo16606	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo35905	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo13929	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo33813	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo8337	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo9487	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo17489	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo19398	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo48945	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo16329	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo1498	Firmicutes	Clostridia	Clostridiales	Veillonellaceae
denovo29110	Firmicutes	Clostridia	Clostridiales	Veillonellaceae
denovo29427	Firmicutes	Clostridia	Clostridiales	Veillonellaceae
denovo45704	Firmicutes	Clostridia	Clostridiales	Veillonellaceae

TABLE 18. Core OTUs identified in 95% of the fecal samples sequenced with the Illumina MiSeq.

MiSeq Fecal Samples Core Biome				
OTU	Phylum	Class	Order	Family
denovo674	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo2755	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo4599	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo5150	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo5489	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo9785	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo23221	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo25947	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo29802	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo50615	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo51252	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo51632	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo53752	Bacteroidetes	Bacteroidia	Bacteroidales	unknown

denovo59784	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo17842	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo32127	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo42471	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo2034	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo56015	Bacteroidetes	Bacteroidia	Bacteroidales	BS11
denovo37440	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo55646	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo13336	Bacteroidetes	Bacteroidia	Bacteroidales	RF16
denovo19430	Cyanobacteria	4C0d-d	YS2	unknown
denovo21705	Cyanobacteria	4C0d-d	YS2	unknown
denovo32398	Fibrobacteres	Fibrobacteria	Fibrobacterales	Fibrobacteraceae
denovo6217	Firmicutes	Clostridia	Clostridiales	unknown
denovo7240	Firmicutes	Clostridia	Clostridiales	unknown
denovo21026	Firmicutes	Clostridia	Clostridiales	unknown
denovo26006	Firmicutes	Clostridia	Clostridiales	unknown
denovo35946	Firmicutes	Clostridia	Clostridiales	unknown
denovo36912	Firmicutes	Clostridia	Clostridiales	unknown
denovo52477	Firmicutes	Clostridia	Clostridiales	unknown
denovo52858	Firmicutes	Clostridia	Clostridiales	unknown
denovo53524	Firmicutes	Clostridia	Clostridiales	unknown
denovo11243	Firmicutes	Clostridia	Clostridiales	Mogibacteriaceae
denovo38012	Firmicutes	Clostridia	Clostridiales	Mogibacteriaceae
denovo41931	Firmicutes	Clostridia	Clostridiales	Mogibacteriaceae
denovo50892	Firmicutes	Clostridia	Clostridiales	Mogibacteriaceae
denovo45338	Firmicutes	Clostridia	Clostridiales	Clostridiaceae
denovo6455	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae
denovo8196	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae
denovo13471	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae
denovo17895	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae
denovo35479	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae
denovo36528	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae
denovo37287	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae
denovo45810	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae
denovo50173	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae
denovo53123	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae
denovo349	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae
denovo55	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo1190	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo1976	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo2834	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae

denovo4922	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo9218	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo10773	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo14867	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo17396	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo17940	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo26375	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo26921	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo29988	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo30284	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo30292	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo31002	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo33893	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo37622	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo38261	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo39321	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo40008	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo46798	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo48169	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo49441	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo49581	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo51969	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo52045	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo52122	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo52848	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo53415	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo55468	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo55929	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo57041	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo57084	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo58053	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo58339	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo60135	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo62074	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo28290	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo32956	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo25139	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo39226	Firmicutes	Clostridia	Clostridiales	Veillonellaceae
denovo19959	Proteobacteria	Alphaproteobacteria	unknown	unknown
denovo25226	Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae
denovo9111	Spirochaetes	Spirochaetes	Spirochaetales	Spirochaetaceae

denovo17103	Spirochaetes	Spirochaetes	Spirochaetales	Spirochaetaceae
denovo24141	Spirochaetes	Spirochaetes	Spirochaetales	Spirochaetaceae
denovo34619	Spirochaetes	Spirochaetes	Spirochaetales	Spirochaetaceae
denovo41824	TM7	TM7-3	CW040	F16
denovo48611	TM7	TM7-3	CW040	F16
denovo6336	Verrucomicrobia	Verruco-5	WCHB1-41	RFP12
denovo6411	Verrucomicrobia	Verruco-5	WCHB1-41	RFP12
denovo22356	Verrucomicrobia	Verruco-5	WCHB1-41	RFP12
denovo54158	Verrucomicrobia	Verruco-5	WCHB1-41	RFP12

TABLE 19. Core OTUs identified in 95% of the cecal samples sequenced with the Roche 454.

454 Cecal Core Biome				
OTUs	Phylum	Class	Order	Family
denovo1039	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo4796	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo6249	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo10009	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo11759	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo21745	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo22494	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo25662	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo27345	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo33845	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo34422	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo35098	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo36633	Bacteroidetes	Bacteroidia	Bacteroidales	unknown
denovo26100	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo30026	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo38089	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo2815	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo15761	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo36873	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo1171	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo15050	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo16036	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo25449	Bacteroidetes	Bacteroidia	Bacteroidales	Paraprevotellaceae
denovo2096	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo4746	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo5091	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo12884	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae

denovo14135	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo15713	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo17690	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo25129	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo26193	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo32707	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo9540	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae
denovo11106	Bacteroidetes	Bacteroidia	Bacteroidales	RF16
denovo21888	Bacteroidetes	Bacteroidia	Bacteroidales	RF16
denovo32221	Bacteroidetes	Bacteroidia	Bacteroidales	RF16
denovo24090	Firmicutes	Clostridia	Clostridiales	unknown
denovo25690	Firmicutes	Clostridia	Clostridiales	unknown
denovo22789	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae
denovo11101	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae
denovo29734	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae
denovo2127	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo9289	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo26711	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo21681	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo14025	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae
denovo9515	Firmicutes	Clostridia	Clostridiales	Veillonellaceae
denovo22931	Firmicutes	Clostridia	Clostridiales	Veillonellaceae
denovo1287	Firmicutes	Clostridia	Clostridiales	Veillonellaceae
denovo9634	Firmicutes	Clostridia	Clostridiales	Veillonellaceae
denovo36177	Firmicutes	Clostridia	Clostridiales	Veillonellaceae
denovo38351	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Verrucomicrobiaceae